

## Duquesne University Duquesne Scholarship Collection

---

Electronic Theses and Dissertations

---

Summer 2007

# Differences in Types and Severity of Consequences of High-Stakes Testing as Realized Between Low and High Performing High Schools and the Impact on the Instructional Programs in Their Schools

Samuel Signorino

Follow this and additional works at: <https://dsc.duq.edu/etd>

---

### Recommended Citation

Signorino, S. (2007). Differences in Types and Severity of Consequences of High-Stakes Testing as Realized Between Low and High Performing High Schools and the Impact on the Instructional Programs in Their Schools (Doctoral dissertation, Duquesne University). Retrieved from <https://dsc.duq.edu/etd/1196>

This Immediate Access is brought to you for free and open access by Duquesne Scholarship Collection. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Duquesne Scholarship Collection. For more information, please contact [phillips@duq.edu](mailto:phillips@duq.edu).

DIFFERENCES IN TYPES AND SEVERITY OF CONSEQUENCES OF HIGH-  
STAKES TESTING AS REALIZED BETWEEN LOW AND HIGH PERFORMING  
HIGH SCHOOLS AND THE IMPACT ON THE INSTRUCTIONAL PROGRAMS IN  
THEIR SCHOOLS

by

Samuel Signorino

Submitted in partial fulfillment of

the requirement for the degree

Doctor of Education

Interdisciplinary Doctoral Program for Educational Leaders

School of Education

Duquesne University

August 4, 2007

**DUQUESNE UNIVERSITY  
SCHOOL OF EDUCATION  
INTERDISCIPLINARY DOCTORAL PROGRAM FOR  
EDUCATIONAL LEADERS**

*Dissertation*

Submitted in Partial Fulfillment of the Requirements  
For the Degree of Doctor of Education (Ed.D.)

*Presented by:*

**Samuel Signorino**  
**B.M. , Music Education, West Virginia University, 1980**  
**M.Ed., Educational Administration, Penn State University, 1988**

*April 19, 2007*

DIFFERENCES IN TYPES AND SEVERITY OF CONSEQUENCES OF HIGH-STAKES  
TESTING AS REALIZED BETWEEN LOW AND HIGH PERFORMING HIGH  
SCHOOLS AND THE IMPACT ON THE INSTRUCTIONAL PROGRAMS IN THEIR  
SCHOOLS

*Approved by:*

\_\_\_\_\_, Chair  
Dr. Phillip Belfiore, Ph.D.  
Professor of Education, Special Education Director, Graduate Education  
Mercyhurst College

\_\_\_\_\_, Member  
Dr. Richard O'Dell, Ph.D.  
Associate Professor of Education  
Mercyhurst College

\_\_\_\_\_, Member  
Dr. Douglas Allen, Ed. D.  
Superintendent, Jamestown School District

**Program Director**  
**James E. Henderson, Ed.D.**  
**Professor of Educational Leadership and**  
**Director, Interdisciplinary Doctoral Program for Educational Leaders**  
**Duquesne University School of Education**

COPYRIGHT

by

Samuel Signorino

2007

## **Abstract**

The purpose of this study was to investigate the role of the Pennsylvania System of School Assessment (PSSA) on the beliefs of teachers and principals, and the impact on the instructional program in their schools. This study collected data on (a) teachers' perceptions of the PSSA, (b) how the program affected instructional practices, and to what extent teachers changed their instructional practices, and (c) what factors, if any, may have influenced these changes. Certified teachers and principals from three high performing and three low performing secondary (9-12 grade) schools in northwestern Pennsylvania were included in this study. Teachers volunteering to participate completed and returned the 54-item, three-part, Likert-style survey indicating to what extent he or she may have changed or altered the use of instructional strategies and techniques used in their classroom, as well as what factors influenced changes in educational practices. Results indicate that the use instructional strategies have changed since the implementation of the PSSA high stakes exams. Contrary to findings in the literature that testing will narrow the curriculum and force teachers to drill and practice the identified content, teachers in all six of the participating schools have increased the use of instructional strategies and tools, considered to be consistent with development of high order skills and best practice, and a decline in use of six items that do not properly involve students in learning. These results lead one to believe that the PSSA assessments did contribute to the changes to instructional classroom strategies as implemented by teachers.

## UMI Abstract

### Abstract

DIFFERENCES IN TYPES AND SEVERITY OF CONSEQUENCES OF HIGH-STAKES TESTING AS REALIZED BETWEEN LOW AND HIGH PERFORMING HIGH SCHOOLS AND THE IMPACT ON THE INSTRUCTIONAL PROGRAMS IN THEIR SCHOOLS

Samuel Signorino

Doctor of Education, August 2007

Duquesne University

Chair: Phil Belfiore, Ph. D.

The purpose of this study was to investigate the beliefs of teachers and principals in regard to the Pennsylvania System of School Assessment (PSSA) and the impact on the instructional program in their schools. This study collected data on (a) teachers' perceptions of the PSSA, (b) how the program affected instructional practices, and to what extent teachers changed their instructional practices, and (c) what factors, if any, may have influenced these changes. Certified teachers and principals from three high performing and three low performing secondary (9-12 grade) schools in northwestern Pennsylvania were included in this study. Teachers volunteering to participate completed and returned the 54-item, three-part, Likert-style survey indicating to what extent he or she may have changed or altered the use of instructional strategies and techniques used in their classroom, as well as what factors influenced changes in educational practices.

Results indicate that the use instructional strategies have changed since the implementation of the PSSA high stakes exams. Contrary to findings in the literature that testing will narrow the curriculum and force teachers to drill and practice the identified content, teachers in all six of the participating schools have increased the use of instructional strategies and tools, considered to be consistent with development of high order skills and best practice, and a decline in use of six items that do not properly

involve students in learning. These results lead one to believe that the PSSA assessments did contribute to the changes to instructional classroom strategies as implemented by teachers.

## TABLE OF CONTENTS

<i>Heading</i>	<i>Page</i>
<b>CHAPTER I: INTRODUCTION .....</b>	<b>1</b>
INTRODUCTION.....	1
HISTORY OF STANDARDIZED TESTING .....	2
STATEMENT OF THE PROBLEM .....	5
NEED FOR THE STUDY .....	12
RESEARCH QUESTIONS.....	13
<b>CHAPTER II: LITERATURE REVIEW.....</b>	<b>14</b>
INTRODUCTION.....	14
TEACHER ATTITUDES.....	14
TEACHER RESPONSE THROUGH CLASSROOM PRACTICE.....	15
STUDENT ATTITUDES .....	19
PRINCIPAL PERCEPTIONS.....	20
SUMMARY .....	21
<b>CHAPTER III: METHODOLOGY .....</b>	<b>23</b>
SAMPLE.....	23
SURVEY.....	24
INTERVIEWS .....	26
FIELD NOTES .....	26
DATA ANALYSIS .....	26
<b>CHAPTER IV: RESULTS .....</b>	<b>29</b>



## TABLE OF CONTENTS (cont.)

INTRODUCTION.....	29
IMPACT ON INSTRUCTION AND INSTRUCTIONAL STRATEGIES.....	36
INSTRUCTIONAL STRATEGIES IN HIGH AND LOW PERFORMING SCHOOLS .....	40
CURRICULUM CHANGES.....	51
PRINCIPAL PERCEPTIONS.....	52
INFLUENCE ON CHANGES .....	53
INSTRUCTIONAL PRACTICES AND YEARS OF EXPERIENCE.....	55
INSTRUCTIONAL PRACTICES AND LEVEL OF EDUCATION .....	57
INFLUENCE.....	58
<b>CHAPTER V: DISCUSSION.....</b>	<b>61</b>
INTRODUCTION.....	61
SUMMARY OF THE STUDY .....	61
METHODOLOGY .....	64
CHANGES TO INSTRUCTIONAL PRACTICES (RQ-1) .....	68
EFFECT OF TESTING ON INSTRUCTIONAL STRATEGIES (RQ-2) .....	69
EFFECT ON CURRICULUM (RQ-3).....	71
PRINCIPAL’S VIEW OF PSSA (RQ-4).....	72
LIMITATIONS.....	76
RECOMMENDATIONS FOR FURTHER RESEARCH .....	77
CONCLUSIONS .....	78

**TABLE OF CONTENTS (cont.)**

**References.....80**

**APPENDIX A: Teacher Survey Instrument.....83**

**APPENDIX B: Principal Interview Questions.....87**

**APPENDIX C: Permission.....90**

## LIST OF TABLES

<i>Table</i>	<i>Page</i>
Table 1: Survey sample.....	30
Table 2: Gender.....	32
Table 3: Principal demographics.....	33
Table 4: Years of experience.....	34
Table 5: Level of education.....	34
Table 6: Area of certification.....	35
Table 7: Increased instructional strategies.....	38
Table 8: Same or decreased instructional strategies.....	39
Table 9: High / Low mean differential.....	41
Table 10: Significant difference of means 2-sample t-test.....	42
Table 11: High-Low chi-square.....	43
Table 12: Chi-square test: High vs. low, no change or decrease.....	45
Table 13: High-Low curriculum.....	47
Table 14: Math and English teachers: Mean responses.....	49
Table 15: Significant difference 2 sample T test.....	50
Table 16: Influence factors.....	54
Table 17: Years experience vs. change.....	56
Table 18: Years experience vs. instructional strategies.....	56
Table 19: Educational level vs. Part I, Part II.....	58
Table 20: Years of experience vs. influence for change.....	59
Table 21: Years experience compared to influence for change.....	60

## CHAPTER I: INTRODUCTION

### Introduction

Over the last twenty years the importance and use of standardized testing in public schools has increased nationwide. The increased pressure from politicians, the media, and the public to create a system of accountability and increased levels of proficiency was reinforced by the release of *A Nation at Risk* (National Commission, 1983). According to the National Commission (1983), for the nation to remain successful and maintain the ability to compete with others around the world, students must be forced to meet rigorous and measurable standards. These standards would raise the level of expectation and minimum competencies for students in 9-12 grade high schools in the United States.

Test-based accountability provides the basis for the more recent No Child Left Behind (NCLB) legislation of 2002, built around a framework that education in the public schools can be improved and strengthened by employing the strategy of testing all students, rewarding high achieving schools, and sanctioning the low performing schools. According to Jacob (2001), this strategy will encourage students to increase achievement and at the same time cause schools to align curriculum and methods.

A large majority of schools, school districts, and states are now using norm referenced standardized tests to measure student achievement. These high-stakes tests are described as multiple-choice exams that measure the retention of small bits of factual knowledge (Elford, 2002) that serve “as a decision point for some concrete outcome in

life” (Wilde, 2002, p. 35), and permit generalization for larger groups of students (Green, 1991; Elford, 2002).

The standardized test is normally “...administered under uniform conditions and is scored according to well-defined scoring procedures” (Elford, 2002, p. 19), usually has a broader content focus, is normed for interpretation and results, and attempts to insure uniformity (Green, 1991). Since their inception and initial use in schools, standardized tests have been designed to collect information that would allow for comparisons of groups of students across the nation (Popham, 2001).

Research on the consequences of high-stakes testing is limited. Interest for this study is in the differences in types and severity of consequences realized between lower performing high schools and high performing high schools as reported by the Pennsylvania Department of Education and to investigate the beliefs of teachers and principals in regard to the Pennsylvania System of School Assessment (PSSA) exam and the impact on the instructional strategies in their schools. This study will investigate the effects of the results of standardized testing on curriculum and classroom strategies.

### History of Standardized Testing

The history of the development of wide-scale standardized testing in public education is, to say the least, interesting. Standardized testing dates back to the early 1900s; however, it only began to gain attention from the United States government in the late 1950s. Prior to this, the public measured the quality of schools based on information related to funding, physical plant, teachers, and demographics of the students (Haney, Madaus, and Lyons, 1993). In an attempt to create a selection process for acceptance into

higher education, the National Merit Scholarship Corporation was founded in 1955. The tests created by this organization were used to identify high ability students, demonstrating the potential to succeed in college. In the same year, 1955, optical scanning equipment was developed at the University of Iowa. This state-of-the-art equipment provided for fast and cost-effective scoring of tests (Haney et al., 1993).

The federal government became involved in testing public school students as early as 1958 with the enactment of the National Defense Act. This act justified support for testing in the schools for reasons of national defense through the development of mental resources, and provided for the first time, funding for testing to local school districts (Haney et al., 1993). In 1963, Francis Keppel, U.S. Commissioner of Education, with the assistance of private funding, created the National Assessment of Educational Progress (NAEP). This assessment was designed to measure the education level of the United States' K-12 population. This program was later funded entirely by the federal government, becoming the first federally mandated collection of educational data.

The 1964 report, Equality of Education Opportunity, began a shift in how the public judged schools. This report put to rest notions that the condition of the physical plant and resources had any influence on achievement. The public summarized the report as saying, "Schools don't make a difference" (Haney et al., 1993, p. 151). Haney et al. (1993), refers to the following quote from Coleman, Campbell, Hobson, McPartland, Mood, Weinfield, and York, (1966, p. 53), that summarizes the report:

"Schools bring little influence to bear on a child's achievement that is independent of his background and general social context; ...this very lack of an independent effect means that the inequalities imposed on children by their home, neighborhood, and peer environment are carried along to become the inequalities with which they confront adult life at the

end of school. For equality of educational opportunity must imply a strong effect of schools that is independent of the child's immediate social environment, and that strong independent effect is not present in American schools" (Haney, et al., 1993, p. 150-151).

The government, with the implementation of the Civil Rights Act of 1964, bolstered the importance of testing in schools. This act, directed at nondiscrimination, focused attention on "...educational outcomes as measured by tests" (Haney et al., 1993, p. 136). The Elementary and Secondary Education Act (ESEA) of 1965 provided assistance to schools that could document enrollment data showing higher concentrations of low-income families. Attention was centered on disparities in scores between the disadvantaged students compared to the more affluent students. Politicians responded by mandating the elimination of disparities through testing. To measure educational achievement the ESEA (1965) called for annual testing (Haney et al., 1993). ESEA reporting procedures eventually led to the standardization of testing in schools. By 1975, more than 90 percent of the schools in the nation were using the same norm-referenced exam. This testing of basic skills was to be administered periodically and reported by grade level and school performance (Haney et al., 1993). The demand for standardized testing grew, encouraged by industry and other groups unhappy with the performance of graduates (Green, 1991). State testing also grew out of the ESEA. The act provided funding to states to develop testing of the basic skills of elementary and secondary students (Haney et al., 1993). In 1981, the Education Consolidation and Improvement Act called for an evaluation of programs based on nationally normed exams (Haney et al., 1993).

The Augustus F. Hawkins-Robert T. Stafford Elementary and Secondary School Improvement Amendments of 1988 approved the development and implementation of testing public and private students in 11th grade to identify high achieving students (Haney et al., 1993).

Other legislation that promoted the use of standardized testing includes the Education for all Handicapped Children Act (EHCA) of 1975, which called for tests to be used for placement, assessment of need, and success of the educational program (Haney et al., 1993). The Individuals with Disabilities Education Act (IDEA) of 1990 (formerly the EHCA) permitted the use of standardized tests for “nondiscriminatory multidisciplinary assessment” (Haney et al., 1993, p. 143). In 2001, President Bush championed the passing of the No Child Left Behind Act. This legislation requires states to create a system of rewards as well as sanctions for schools based strictly on student and or school performance (Nitko, 2004). It also requires schools to demonstrate proficiency in reading and math of all students by the year 2014.

#### Statement of the Problem

In an attempt to hold students, teachers, and schools accountable, a large majority of schools, school districts, and states have elected to use norm-referenced standardized tests to measure student achievement. Moon, Brighton, and Callahan (2002) report 49 states have allocated funding for the development of testing instruments to assess effectiveness of schools. The design of standardized, norm-referenced tests is intended to provide educators with information about how well the students are achieving in academic areas, as compared to similar groups of students (Kohn, 2000).



This study grew out of the resulting consequences of high-stakes testing. The researcher was interested in (a) the differences in types of consequences realized between lower and higher performing high schools as reported by the Pennsylvania Department of Education, (b) investigating the beliefs of teachers and principals in regard to the Pennsylvania System of School Assessment (PSSA) exam, and (c) the impact of PSSA results on the instructional programs in their schools.

The business community and government officials have been calling for improved levels of achievement from our students, resulting in increased testing in our schools. It is possible that these same groups will eventually realize that increased use of a single test, used for decisions concerning students, curriculum, and schools, will not produce the long-sought results of a highly educated and capable society, but instead a narrowed and specific curriculum that will make it possible for schools to show that every student is capable of scoring proficiently on the exam. Schools may be graduating students with an extremely high ability to recall data as related to specific factual knowledge, students with a lack of understanding as to why or how things and concepts have been developed, and students with a weakened ability to problem solve.

Kohn (1999) suggests that the emphasis currently being placed on test scores will result in increased importance on classroom strategies that focus on the development of basic skills and drill and practice, that standardized tests are a tool used to realign the curriculum to the state-mandated curriculum. However, high-stakes testing has resulted in a drastic change in the way instruction is delivered, and few if any of the changes are positive. It is generally accepted in many schools and communities across the country

that it is acceptable to provide drills and practice for a high-stakes test, even if the students are learning very little (Kohn, 2000).

Teachers may stray from the school district approved curriculum to teach items that will increase scores on tests (Kohn, 2000). Students may be refused enrichment opportunities to make way for practice on the upcoming standardized assessment. Teachers teach skills necessary for test taking rather than the curriculum-based content of the course (Kohn, 1999). As of 1989, each student in the United States was taking two and a half standardized tests per year on average (Green, 1991). Days or weeks, and sometimes months, of instructional time are used to increase test scores. Portions of the various curriculums will be skimmed while other areas that might be tested are emphasized. Practice tests replace guided practice (Kohn, 2000). Classroom instruction often takes on the same format as the test. Not only are the tests and quizzes taking the form of multiple-choice items, but the activities associated with learning also become multiple-choice. Essay tests are sometimes abandoned because they give no advantage to the student taking the standardized test (Kohn, 1999). Green (1991) foresees state-mandated tests for minimum competencies eventually forcing a shift from the local control of curriculum by the school board to control by the state. She explains that high-stakes testing assumes control of the curriculum. The exam determines what will be included and excluded from the curriculum. Eventually, the teaching and learning process is narrowed to the content of the test. Attention to specific skills necessary to score higher on a test may avert focus from more difficult skills such as critical thinking and problem solving (Green, 1991).

Standardized tests are used for selection of individuals, placement of students, planning of instruction, diagnosis, academic counseling, program evaluation, and individual student comparisons to school, district, state, and national norms (Green, 1991). They are also used to demonstrate accountability, to determine priorities in funding, to assess teacher effectiveness, and to certify completion of a program (Green, 1991).

Critics of testing complain that

“Standardized tests can’t measure initiative, creativity, imagination, conceptual thinking, curiosity, effort, irony, judgment, commitment, nuance, good will, ethical reflection, or a host of other valuable dispositions and attributes. Social and moral development will be ignored. Skills such as conflict-resolution, building a sense of community, and allowing for creative play will not be addressed. What these tests can measure and count are isolated skills, specific facts and functions, the least interesting, and least significant aspects of learning” (Kohn, 2000, p. 82).

These tests do not take into account the process used by the test taker to arrive at any given answer, therefore ignoring the question of understanding. Students may select the correct answer and not understand the problem (Kohn, 2000).

Elford (2002) cites the writing of Jacques Barzun (1991), who criticizes standardized testing as a measure of passive recognition knowledge. Elford (2002) also points out the inability of some very capable individuals to successfully take standardized tests and attributes undocumented accounts of less than favorable results on these exams to stress as associated with timed testing. Delisle (1997) refers to

“...countless cases of magnificent student writers whose work was labeled as ‘not proficient’ because it did not follow the step-by-step sequence of what the test scorers (many of whom are not educators, by the way) think good expository writing should look like. And, with many of the multiple choice questions having several ‘correct’ options in the eyes

of creative thinkers, scores get depressed for children who see possibilities that are only visible to those with open minds” (p. 42).

Reporting results of testing in schools often receives negative reviews. Kohn (2000) attributes norm-referenced tests and scoring practices to a feeling of superiority among successful students. He explains that scores from norm-referenced tests never change, that there will always be a top 10 percent. As median indicates the middle, results will always show one half above and one half below the median. Additionally, these scores never indicate the number of items scored correctly or incorrectly. High-achieving students most likely understand that their scores will remain in the higher percentage groupings, and therefore will only work to maintain instead of striving for excellence (Kohn, 2000).

On the other hand, Kohn (2000) reminds the reader that these types of scores, when used to compare students and schools, will always have some group of students scoring low, and because certain schools have a larger proportion of this group, those schools will appear to be failing for this same reason. The general public does not always understand the reason for testing, nor do they understand the reporting of the results. Test marketing has created an image that is misleading to the consumer (Kohn, 2000).

Other possible consequences of high-stakes testing include the narrowing of instruction; weakening of pedagogy; declining innovations; a shift from student-centered to teacher-centered classrooms (Saltman & Gabbard, 2003); unhealthy, unethical competition among teachers; coaching students during exams; changing answers on the exam (Wilde, 2004; Kohn, 2000); alienation of student groups; grade-level retention; and teaching to the test (Wilde, 2000). Teachers may begin to dislike having the lower level

student in their classes because they might reduce the perceived success of the teacher (Kohn, 2000).

It is possible that schools will lose the locally created, meaningful curriculum to the state or national-normed, test-oriented instruction and will soon begin to suffer the loss of highly qualified teachers and principals. Kohn explains that teachers tire of having to teach a test-driven curriculum, and principals are reluctant to take on a no-win situation of raise the scores. Many will simply leave. Over time, this loss of experienced professional staff may actually reduce the standard of education (Kohn, 2000). Increased use of high-stakes testing and pressure on the teacher to deliver content may increase the need for specialization in certification. This, in turn, may spread the departmentalization and ability groupings currently seen in high schools into the lower elementary grades (Kohn, 2000).

Standardized testing can produce useful information when used as intended. The design of standardized, norm-referenced tests is to provide educators with information about how well the students are achieving in academic areas, as compared to similar groups of students. However, with the current use of standardized, norm-referenced tests, individuals and schools are to be accountable for factors that are beyond their power to control. Low scores are often the result of the socioeconomic surroundings of the student, which include the school resources as well as the affluence of the school community. Schools are forced to test frequently in order to increase familiarity of teachers to standardized tests. Schools are provided monetary incentives to pay attention to test scores, such as bonuses for high scores and penalties such as reduction of revenue or loss of graduation (Kohn, 2000).

Individual school and district scores from across the nation are published annually, comparing schools and placing pressure on educators to increase scores (Kohn, 2000). Widespread concern exists in the educational community regarding the exaggerated importance of one single assessment score and unintended consequences (Wilde, 2002), and the use of test scores as the only measure of ability brought about by calls for accountability. This creates the sense of a special circumstance, causing teachers and students to produce an out-of-the-ordinary performance response (Elford, 2002).

The Pennsylvania Department of Education has implemented its own version of high-stakes testing currently known as the Pennsylvania System of School Assessment (PSSA). Students are required to score at a state-mandated level of proficiency in reading, and mathematics before becoming eligible to receive a high school diploma. And more recently, the federal government, through NCLB legislation, has required schools, by the year 2014, to have all students proficient in reading, and mathematics. With this mandate also comes an increased amount of testing for students (National Association of Elementary School Principals & National Association of Secondary School Principals, 2003).

One must agree that the single assessment programs such as those mandated at the state level are lacking in a number of ways. The tests do not provide the most accurate representations of student understanding, achievement, ability levels, or performance levels. There is no guarantee that the test matches the locally selected curriculum, and there is even less promise that the test will actually include items that measure the important teacher-stressed content.

## Need for the Study

Little research has been initiated concerning the consequences of high-stakes tests. Recent studies (Moore, 1994; Moon et al., 2002; Hoffman, Assaf, and Parris, 2001) point to an increased interest in the effect of the test on the student, the teacher, and the curriculum. It is important to note that of the available research concerning state-mandated or high-stakes testing, almost all of it has been conducted at the elementary school level (e.g. Moore, 1994; Moon et al., 2002). One reason for this could be that the researchers had easier access to elementary schools, or that elementary teachers were or are more vocal about the use of standardized testing. It may also be possible that school districts, having already received criticism about test scores, are reluctant to promote further scrutiny about test scores at the secondary level. This, however, is where the high-stakes testing has been aimed. Business and industry, as well as the politicians and now the general public, are all watching to see the results of the increased emphasis on testing for proficiency. These groups are hoping to see students graduating from our nation's high schools with higher levels of knowledge, increased use of technology, and most important, creativity, initiative, and problem-solving skills that graduates of years past may not have achieved. They are also watching for schools to meet the requirement of the No Child Left Behind (NCLB) Act of 2001, that all students (100 percent) "...be performing at or above proficient levels in reading and mathematics by the end of the 2013-2014 school year" (National Association of Elementary School Principals & National Association of Secondary School Principals, 2003, p. 17), with accountability measured by producing proficiency at the 100 percent level.

## Research Questions

1. How does the PSSA math and reading assessments, since the reporting of scores and institution of NCLB, impact instruction and instructional strategies?
2. How do high-stakes testing results (PSSA Math and Reading) and the requirements of NCLB affect instructional strategies in high and low performing high schools?
3. How do high-stakes testing results (PSSA Math and Reading) affect curriculum in high and low performing high schools?
4. How do principals in high and low performing high schools view the PSSA standardized testing and accountability program?
5. What factors have influenced these changes in high and low performing high schools?



## CHAPTER II: LITERATURE REVIEW

### Introduction

This chapter will review the literature on high-stakes testing, beginning with the perceptions and attitudes of teachers toward high-stakes standardized testing, followed by a discussion about how teachers use results of these exams to modify instruction and classroom strategies. The third section will review the attitudes and beliefs expressed by students. The last sections of this chapter will review attitudes of principals toward high-stakes testing.

### Teacher Attitudes

The study by Moore (1994) explored teacher attitudes toward standardized testing and found that teachers reported the standardized testing program was of no value to them and that it did not improve student learning. He concluded that teacher perceptions about standardized testing programs were not positive. Hoffman et. al (2001) reported dissatisfaction with the standardized test and the process on the part of the teachers. Teachers claimed that the test (a) administration requires too much time to be taken from the normal curriculum, (b) has gained too much importance and is pushed by administration, and (c) is not worth the cost and time and should be eliminated. Ninety percent of the teachers participating in the study by Edelman (2001) indicated that the cost of the testing far outweighed the benefits received. Jones et al. (1999) reported a belief by teachers that low morale was and will continue to be a result of the mandated testing program, with pressure to keep good scores high or to increase scores from previous testing as the reason. However, when offered a course in tests and

measurements, and subsequently understanding the purpose and use of the test, teachers showed a gain in positive attitude toward standardized testing (Green & Eicher, 1987). Jones et al. (1999) also indicate that teachers (61 percent) believe that students display a higher level of anxiety with standardized testing, and (48 percent) that the testing program negatively affects the students' thirst for knowledge.

Abrams et al. (2003), reports on a survey conducted by the National Board on Educational Testing and Public Policy, examining teachers' opinions and attitudes about state testing programs. Teachers were asked to complete an 80-item Likert-style survey with questions or statements focused on the impact of testing in the classroom and pressure to improve scores and test preparation. Results of the survey indicate that just over half (58 percent) of the respondents agree that the state tests are based on a curriculum that should be followed. They also agree that if the standards are taught properly, the students will do well on the exam.

#### Teacher Response through Classroom Practice

Whereas research on students' attitudes toward high-stakes testing is minimal, research on how teachers use standardized testing data to adjust curriculum is more numerous. Teachers responding to a survey by Jones et al. (1999) and affirmed by Abrams et al. (2003) indicated that time spent on instruction has changed and that instructional strategies have been altered since the implementation of standardized testing. Vogler (2002) found similar results, reporting that teachers changed their instructional practices by lessening the importance on strategies that promote higher order thinking skills, while at the same time increasing practices that do not develop these

same skills. Additionally, Moon et al. (2002) investigated the perceptions of students and teachers concerning the effects of high-stakes testing strategies. Subject areas that are tested have been given priority over subjects without a test and are regularly provided larger blocks of time while those subjects without state tests are afforded low priority and reduced time (Abrams et al., 2003; Moon et al., 2002). An interesting finding was that more than 75 percent of the respondents reported that the testing programs have caused them to teach in ways that are not sound educational practices, and that the program negatively impacts the quality of instruction (Abrams et al., 2003).

Moore (1994) developed the Teacher Assessment Practices Questionnaire (TAPQ), consisting of 71 items designed for elementary teachers. Teachers choosing to participate in this study indicated that they used the test results to (a) determine individual and group achievement, (b) report progress to parents, (c) determine educational levels, (d) group pupils, (e) analyze learning problems, and (f) identify the need for corrective work. The teachers also indicated that they used the test results to guide decisions regarding instruction and curriculum by placing greater emphasis on test material already contained in the existing curriculum. Others in the study augmented the regular curriculum to include test-related topics, while some omitted enrichment material in an attempt to spend more time on test preparation. Moon et al. (2002) also found that teachers reported spending large amounts of time preparing students for state-mandated testing and the omission of student-centered, time-intensive, or creative activities. Teachers increased the used of worksheets, test taking strategies, more hands-on activities, practice tests, and drill and practice exercises. Interviews of students and teachers indicated that the lessons focused on test preparation were much of the time

spent simulating testing experiences (Moon et al., 2002; Jones et al., 1999). Edelman (2001) also attempted to determine what impact the mandated state-testing program had on classroom practices. The researchers used a questionnaire, sent to third grade teachers. Respondents in a study by Edelman (2001) indicated that they prepared students for the exams by teaching time management and stressing completion of answer sheets. They also indicated an increased emphasis on subject matter and specific information known to be on the test.

Moore (1994) listed three examples of relationships among perceptions and practices. First, a teacher feeling pressure to improve test scores would be less likely to use the data to identify additional activities for the students. Next, less instructional time will be spent on enrichment activities when greater pressure is exerted to improve test scores. And third, teachers finding the standardized testing program to be of little value or having a poor perception of testing are more likely to participate in the inappropriate practices discussed earlier.

Teachers interviewed expressed a concern that they are forced to offer as much simulation practice as possible, leaving out much of the interesting and beneficial activities, information, and skills normally taught while practicing for the test. Moore (1994) reported a moderate to extreme amount of pressure on the teachers to improve standardized test scores from the principal. Hoffman, Assaf, and Paris (2001) also found this to be true. Participants in their study included 500 randomly selected individuals, all members of the International Reading Association. Teachers were asked to respond to a survey consisting of 113 Likert-style items focused on attitudes, test preparation, administration practices, uses of scores, effects on students, and overall impressions of

the standardized test. Hoffman, Assaf, and Paris (2001) found teachers willing to report unethical practices such as giving hints to students, pointing out miss marked answers, providing instruction during the test, and pointing out incorrect answers. Teachers also reported substantial pressure from administration and parents to improve scores on the test, with 41 percent reporting that they taught only what was on the test, eliminating everything else in the curriculum (Abrams et al., 2003). This suggests that state-mandated testing might cause a narrowing of the curriculum. Teachers reported substantial time spent in test preparation and teaching test-taking skills. Some reported the use of commercially developed materials while others reported the use of state-released items.

A large percentage of the teachers (85 percent of 79 respondents) believe that the test scores and the campaign to improve these scores have gained too much emphasis (Moore, 1994). The teachers think they teach more to the test and usually omit pertinent information for lack of time. Teachers admitted to the use of unethical or improper practices in preparing students for the exams. Examples include the use of the prior year's test as practice, changing of student answers, extending the time beyond the maximum limit, helping students during testing, and providing practice items from the actual test. This could possibly be a reaction to the pressure to improve. Others may be oblivious to the fact that they are using unacceptable means (Moore, 1994). Moon et al. (2002) suggests that these classroom strategies, prioritized scheduling, test-taking lessons, drill and practice, and the removal of enrichment activities may actually cause a negative growth in achievement.

## Student Attitudes

Jones et al. (1999) report the results of statewide surveys administered to 236 elementary teachers in five school districts in North Carolina. The survey questioned teachers on their opinions of the state-mandated accountability program and how their students are affected by the testing. Respondents reported that students exhibited emotional effects that include crying, vomiting, and fighting. Hoffman, Assaf, and Paris (2001) report that many students exhibit stomachaches and headaches while taking standardized exams. They also explain that many students become anxious, irritable, and aggressive while testing.

Research (Kohn, 1999; Wideen, O'Shea, Pye, & Ivany, 1997) shows a decline in the value students place in learning when greater emphasis is placed on grades and test scores. Student attitudes toward learning are reduced drastically when repeatedly reminded and encouraged to assess their own learning or performance (Kohn, 1999).

Eakins, Green, and Bushell (2001) wanted to determine if (a) students that receive practice with an instructional unit will achieve at a higher rate than those without practice, and if (b) students who receive multiple practices show higher achievement than those with only one practice. Participants included 170 first grade students from three inner-city elementary schools in a large metropolitan area. Students were administered the Metropolitan Achievement Test (MAT) as the dependant variable, while the Behavior Analysis Test (BAT) served as the independent variable. The BAT is a unit designed as a tool to assist in the development of test-taking skills. It was presented singularly or repeated many times (at least three), depending on the group of students being tested.

Eakins et al. (2001) found that students scored higher on achievement tests when they received instruction on test-taking strategies. They also report increased gain when students received multiple presentations of the test-taking course.

### Principal Perceptions

A study by McCall (2003) investigated principals' perceptions of the Massachusetts Comprehensive Assessment System. Twelve principals with more than ten years of experience at the elementary/middle school level were interviewed. Contrary to the opinion of the teachers, the participating principals believed that the public release of test scores helps to motivate educators to increase student outcomes. They feel that the purpose of the public release is to force schools to realign the curriculum with set standards.

The accountability created by the testing affects the amount of time and effort that principals place on achievement of increased test scores. McCall (2003) reported principals spending more time focusing on curriculum than before mandated testing was implemented. The participants indicated that the curriculum frameworks, otherwise referred to as standards, took precedence over the state test. Reed, McDonough, Ross, and Robichaux (2001), in a study investigating principals' perceptions of the impact of high-stakes testing, found that principals of higher achieving schools place less pressure on teachers to increase scores on state tests. On the other hand, principals from lower achieving schools are deeply concerned about testing results. Administrators from these schools talk about the enormous impact of high-stakes testing on their schools.

Principals favor the use of test results to adjust curriculum. The standards have helped to create a more consistent learning opportunity because the principals choose to match the curriculum with the state standards. Principals report that they do use test results while developing improvement plans and to realign portions of the curriculum. They also indicate that they would be more likely to address inefficient teaching methods and skills (McCall, 2003).

### Summary

The review of the literature on high-stakes standardized testing indicates a lack of teacher support and negative perceptions (Moore, 1994), dissatisfaction with the process, a belief that valuable instructional time is used for test preparation and administration, and that the testing benefits are not worth the cost (Hoffman et al., 2001). Time spent on instruction has changed and strategies have been altered since the implementation of standardized testing (Jones et al. 1999; Abrams et al. 2003; Vogler, 2002). Tested subject areas have gained priority over other subjects (Abrams et al., 2003; Moon et al., 2002), and elementary teachers have reported that testing programs cause them to resort to less than sound pedagogical practices (Abrams et al., 2003). Testing has caused teachers to spend large amounts of time preparing students for state-mandated testing as well as the omission of student-centered, time-intensive, or creative activities (Moon et al., 2002).

Student perceptions toward standardized testing also tend to be negative (Jones et al., 1999). As greater emphasis is placed on scores, student attitudes toward learning are reduced (Kohn, 1999; Wideen, O'Shea, Pye, & Ivany, 1997).



Principals, however, believe that the reporting of test scores helps to motivate educators to increase student outcomes. They feel that the attention placed on test scores helps to force schools to realign the curriculum with set standards (McCall, 2003).

Principals of higher achieving schools place less pressure on teachers to increase scores on state tests while principals from lower achieving schools are deeply concerned about testing results (Reed, McDonough, Ross, & Robichaux 2001).

### **CHAPTER III: METHODOLOGY**

Little research has been conducted analyzing the resulting consequences of high-stakes tests. The purpose of this study was to investigate the beliefs of teachers and principals in regard to the PSSA, and the impact of the results on the instructional program in their schools. Additionally, this study examined the differences in types and severity of consequences realized between lower performing high schools and high performing high schools as reported by the Pennsylvania Department of Education. This study collected data on (a) how high-stakes testing results affect instructional strategies in high and low performing high schools; (b) how high-stakes testing results affect curriculum in high and low performing high schools; (c) how principals view the PSSA standardized testing and accountability program in high and low performing high schools; and (d) what factors have influenced these changes in high and low performing high schools?

Data were also collected from high school principals concerning their views on these same issues surrounding the PSSA standardized testing program and how it has affected the instructional program in their schools.

#### **Sample**

The study included certified teachers and principals from six secondary (9-12 grade) schools in Pennsylvania. The schools were located in the northwestern region of Pennsylvania. Schools were selected based on the test results from the 2004-2005 school year administration of the PSSA 11th grade reading and math exams as reported by the Pennsylvania Department of Education.

Current reporting practices list the percentage of students per school scoring at the advanced, proficient, basic, and below basic levels of proficiency. Three of the six high schools included in the study were high performing schools, or schools that realized overall scores above the 70 percent mark of students at the proficient or advanced level in both math and reading. The remaining three schools were included based upon overall scores of 35 percent or higher number of students scoring below proficient (basic and below basic) on the latest administration of the PSSA reading and math exams. Actual 2004-2005 PSSA math scores from schools within the region ranged from a low of 31 percent of students proficient or above to a high score of 89.9 percent proficient or above. On the other hand, the percentage of students in the region scoring below proficient on the 2004-2005 PSSA math exam, ranged from a low of 10 percent to as high as 69 percent. Although similar, the 2004-2005 reading scores for the county were slightly higher as compared to the math scores. Proficient or above scores for students in the region ranged from as low as 47.3 percent to a high of 96 percent. Students scoring below proficiency in reading in the region ranged from a low of 1.3 percent to a high score of only 52.8 percent. Schools were purposefully selected based on scores, location, and availability to the researcher.

### Survey

Principals of the included high schools were asked to allow the researcher to address a faculty meeting to distribute a survey to certified teachers, in each of the specialized areas of English, Math, Science, Social Studies, Art, Music, Physical Education, Business, Industrial Arts, Foreign Language, and Special Education. After an

introduction, a brief outline of the study, and time for questioning, participants were given time to read and sign the letter of consent to participate in the study. Teachers volunteering to participate were encouraged to complete the survey and return completed surveys to the researcher before leaving the meeting.

The survey instrument was developed for a study completed by Vogler (2002). The purpose of the Vogler (2002) study was to ascertain what impact, if any, the release of the Massachusetts Comprehensive Assessment System (MCAS) test results had on instructional practices. The instrument required minor adjustment to coordinate with the Pennsylvania System of School Assessment. One particular change was from a reference to the MCAS to the PSSA exam. Another change asked the participant for additional demographic data.

The 54-item, three-part survey asked the respondent to indicate to what extent they may have changed or altered the use of 40 instructional strategies and techniques possibly used in their classroom or school. Section one asked the respondents to indicate a degree of change in usage for each item. Likert-style responses asked the respondent to indicate large increase in use (LD), decrease in use (D), same or steady use (S), increase in use (I), or large increase in use (LI). Respondents were also permitted indicate that the item does not apply to their curricular area (NA).

The second section of the survey asked teachers to respond to 10 items indicating what factors influenced changes in educational practices. This portion of the survey again required Likert-style responses ranging from strongly disagree to strongly agree. And finally in section three of the survey, respondents were then asked to supply demographic information describing sex, years of teaching experience, level of education

or degree, and area of certification (Vogler, 2000). The data was analyzed with the assistance of Excel and Mini Tab statistical software.

### Interviews

Principals of each of the schools were also asked to participate in a face-to-face interview, at a time and place convenient for the participant. After a brief introduction and an outline of the researchers background, a discussion concerning the study, and an opportunity to answer questions, the administrator was provided with the opportunity to read and then sign the letter of consent for participants. The interviews were be taped, when permitted, and then transcribed. Responses were analyzed using a process of induction.

### Field notes

Field notes were taken during each session to allow for recording of observations, thoughts, and insights, through the interview process. These notes were intended to serve as an additional source of data, lending trustworthiness to the study, as well as strengthening the content validity and reliability of the data collection process.

### Data Analysis

The data was analyzed using Excel and Minitab statistical software. Responses to the survey questions in section one and two were analyzed using descriptive explanatory techniques calculating the mean, and mode of the responses (Fink, 1995). This information was then be used to compare results of the three high performing schools to the answers provided by the respondents in the three lower performing

schools. The differences and range of differences between the high and low performing schools demonstrated where the schools actually differ on the items and issues presented in the survey.

Section one and two were also be analyzed by percentage of respondents per survey question allowing the researcher to report the proportion of respondents indicating an increase or decrease in the use of each of the following practices since the initial implementation of the PSSA test results. Excel was used to analyze descriptive statistics and to run the Pearson's correlation.

Data was also used to test for a possible correlation between the differences in mean between the high performing schools and the low performing schools. Section one and two data was also used in chi-square (goodness of fit) tests to compare differences between the high and low performing schools.

An analysis of section three provided demographic data detailing the level of experience and assignment of the teachers responding to the survey. Using these data, the researcher investigated possible correlations between a) use of instructional practices and levels of experience b) use of instructional practices and level of education c) reasons for change as compared to level of education d) reasons for change compared to levels of experience. Having these data as well as the actual PSSA scores from the Pennsylvania department of Education allowed for comparison between scores and individual strategies, scores and changes in practices, scores and levels of experience, and scores and levels of education.

Through the survey of high school teachers and interviews of principals in each of the selected schools, this study investigated the specific consequences of testing our high

school students. The expectation was that the less successful, lower performing high schools are experiencing higher levels of narrowing of instructional strategies. It was also expected that they were experiencing a reduction in enrichment activities throughout the curriculum, and eventually a narrowing of the overall curriculum. On the other hand, highly successful, high performing high schools were not seeing negative effects toward teaching strategies. These schools may not have had to adjust the curriculum to accommodate student scores below the proficient level.

## **CHAPTER IV: RESULTS**

### **Introduction**

This study examined the instructional consequences of high-stakes testing. The researcher was interested in (a) the differences in types of consequences realized between lower and higher performing high schools as reported by the Pennsylvania Department of Education, (b) investigating the beliefs of teachers and principals in regard to the Pennsylvania System of School Assessment (PSSA) exam, and (c) the impact of PSSA results on the instructional programs in their schools. After a brief section outlining the demographics of the participants in the study, the remainder of the chapter is organized in terms of the specific research questions posed in Chapter 1. The first section will present the data showing how the PSSA math and reading assessments, since the reporting of scores and institution of NCLB, impact instruction and instructional strategies. How high-stakes testing results (PSSA Math and Reading) and the requirements of NCLB affect instructional strategies in high and low performing high schools will be addressed in section two. The third section will provide data that shows how high-stakes testing results (PSSA Math and Reading) affect curriculum in high and low performing high schools. The last section will show what other factors have influenced these changes in high and low performing high schools.

### **Demographics of Respondents**

Of the six schools selected to be included in the study, three were high performing schools, realizing scores above the 70 percent mark of students at the proficient or advanced level in both math and reading on the 2005 administration of the 11<sup>th</sup> grade PSSA exam and as reported by the Pennsylvania Department of Education. To maintain



confidentiality, these schools shall be referred to as schools A, B, and C. The remaining three schools were included based upon scores of 35 percent or higher number of students scoring below proficient (basic and below basic) on the same administration of the PSSA reading and math exams during the 2004-2005 school year. For purposes of this study, the low performing schools shall be referred to as X, Y, and Z.

Respondents from the three high performing schools totaled 106 teachers or 84.13% of the 126 potential participants. Thirty-six teachers or 87.8% were from school A, 46 teachers or 93.88% were from school B, and 24 teachers or 66.67% were from school C. Respondents from the three low performing schools totaled 91 teachers or 73.98% of the total 123 potential participants. The principal from each of the six high schools also agreed to participate in the study.

Table 1 shows the number of respondents as compared to the total number of faculty members in each of the six high schools.

Table 1

*Survey Sample*

School	Total Faculty	Number Respondents	Percent Respondents	Participating Principals
A (High)	41	36	87.80%	1
B (High)	49	46	93.88%	1
C (High)	36	24	66.67%	1
X (Low)	40	33	82.50%	1
Y (Low)	38	25	65.79%	1
Z (Low)	45	33	73.33%	1
All Schools	249	197	79.12%	6
High Performing (ABC)	126	106	84.13%	3
Low Performing (XYZ)	123	91	73.98%	3

As can be seen in Table 1, the number of participating faculty members from high performing schools was very close to the number of participating faculty members from the low performing schools. However, the response rate from the high performing school faculty was over 10% higher than those choosing to participate from the low performing schools. Response rates from the three high performing schools ranged from as low as 66.67 % to a high of 93.88 %, with a mean of 84.13 %. The range of response rates from the low performing schools ranged from a low of 65.79 % to a high of 82.5 %, with a mean of 73.98 %. The mean response rate from all six participating high schools equaled 197 of 249 teachers or 73.98 %.

As stated earlier, one principal from each of the high schools agreed to participate in the study. The principals from the 3 high performing schools (A, B, and C) were all male. Two of the principals from the low performing schools were male (E and F). Only the principal from school D was a female.

The teacher respondents included 109 female participants or 55.33% of the sample, and 88 male participants or 44.67% of the sample. The participating teachers from the high performing schools included more males (54) than females (52). On the other hand, female respondents from low performing schools outnumbered the male participants by an amount just under two to one, with female respondents equaling 57 or 63.1 percent and only 34 or 36.9 percent male. Table 2 shows the comparison of male respondents to female respondents from the included schools.

Table 2

*Gender*

Schools	Male	% Male	Female	% Female
A	17	47.2%	19	52.8%
B	24	52.2%	22	47.8%
C	13	54.2%	11	45.8%
X	15	45.5%	18	54.5%
Y	8	32.0%	17	68.0%
Z	11	33.3%	22	66.7%
Totals	88	44.67%	109	55.33%

Each participant in the study, including principals, reported years of experience. The three principals from the high performing schools (principals A, B and C) reported eight years of teaching experience each. Experience in the classroom varied in the low performing schools with principal D reporting 18 years, principal E reporting 11 years, and principal F reporting 5.5 years of teaching. Administrative experience as well as other demographic information about each principal can be found in Table 3.

Table 3

*Principal Demographics*

---

Principal	Gender	Teaching Experience	Administrative Experience	Principal of current school	Teaching Certification
A	Male	8 years	18 years	9 years	Music Education
B	Male	8 years	18 years	10 years	Music Education
C	Male	8 years	9 years	4 years	Mathematics
X	Female	18 years	7 years	1.5 years	Physically/Mentally Handicapped
Y	Male	11 years	3 years	2 years	Social Studies
Z	Male	5.5 years	3 years	3 years	Science

---

Only 21 of the teachers (10.66%) have been teaching for three years or less. The largest group of teachers, 63 or 31.98% has been teaching 4 to 12 years in the classroom. Thirty-six or 18.27% of the respondents indicated that they had been teaching between 20 and 27 years. Another 37, or 18.78% of the respondents indicated 28 or more years of service. Table 4 shows the number of participating teachers and their total years of experience in the classroom.

Table 4

*Years of Experience*

Years Experience	High Performing	Low Performing	Total	%
3 or less	6	15	21	10.66%
4-12	26	37	63	31.98%
13-19	18	22	40	20.30%
20-27	25	11	36	18.27%
28 or more	31	6	37	18.78%
Totals	106	91	197	100.00%

The respondents also indicated a wide range of educational levels. All of the six participating principals reported completion of a master's degree. The majority of the teachers responding to the survey (104) have earned a bachelor degree. One teacher indicated having completed a doctorate, and the balance (92) has completed the master degree. Table 5 illustrates the level of education achieved by the teachers responding to the survey.

Table 5

*Level of Education*

Degree	Credits	High Performing Respondents		Low Performing Respondents		Total Respondents	
		Number	%	Number	%	Number	%
Bachelor		8	7.55%	19	20.88%	27	13.71%
Bachelor	+15	39	36.79%	38	41.76%	77	39.09%
Master		28	26.42%	20	21.98%	48	24.37%
Master	+15	9	8.49%	9	9.89%	18	9.14%
Master	+30	7	6.60%	3	3.30%	10	5.08%
Master	+45	9	8.49%	1	1.10%	10	5.08%
Master	+60	6	5.66%	0	0 %	6	3.05%
Doctorate		0	0 %	1	1.10%	1	0.51%

Participants in the study also indicated their most current teaching assignment. English teachers totaled 39 (19.8%), 35 (17.77%) in math, 28 (14.21%) in social studies, and 25 (12.69%) in science. Certification in special education totaled 19 (9.64%) teachers, seventeen (8.63%) in physical education, 11 (5.58%) in foreign language, eight (4.06 %) in art, and eight (4.06%) music. The balances of the teachers were assigned in the areas of business, technology education, and agriculture education. Table 6 shows the number of teacher respondents from each of curricular area or teaching assignment.

Table 6

Area of Certification

Department	Number of Respondents	Percentage of All Respondents
English	39	19.80%
Math	35	17.77%
Science	25	12.69%
Social Studies	28	14.21%
Art	8	4.06%
Music	8	4.06%
PE	17	8.63%
Business	3	1.52%
Tech Education	3	1.52%
Special Education.	19	9.64%
Foreign Language	11	5.58%
Other	1	0.51%
English	39	19.80%
Total Respondents	197	100%

## Impact on Instruction and Instructional Strategies

Responses to Part I of the survey instrument can be useful in answering the first research question, how the PSSA math and reading assessments, since the reporting of scores and institution of NCLB, impacted upon instruction and instructional strategies?

Participating teachers responded to 40 separate items focused on instructional strategies, materials, and tools. The responses indicated the extent to which the teacher decreased or increased the use of each of the instructional strategies since the implementation of PSSA and NCLB. Respondents used a Likert style scale allowing for responses indicating a large decrease, decrease, no change, increase, large increase, or not applicable. An increase or decrease in use of each item was determined by finding the mean of the responses for each item. A mean greater than 3.0 ( $>3.0$ ) demonstrated an increase in the use of the instructional strategy, material, or tool by the teachers, while a mean less than 3.0 ( $<3.0$ ) indicated a decrease in those strategies, materials, and tools. The percentage of increase or decrease was determined by calculating the percentage of teacher indicating the increase or decrease of each item.

Of the forty items in Part I of the survey instrument, 36 of the items (table 6) demonstrated a mean greater than 3.0 or an increase of use by the teachers since the implementation of the PSSA and NCLB. As would be expected, the farther the number is away from the 3.0, the larger the change, either positive or negative. The strategy indicated as having the largest increase of use was writing assignments with a total percentage increase of use equal to 70%. Other items showing high percentages of increase of use included rubrics or scoring guides at 63.68%, computer/educational software at 59.34%, open response questions at 58.73%, computers / online research at

58.24%, and creative/critical thinking questions showing an increase in use of 57.22%.

See table 7 for a complete list of items showing an increase of use by participating teachers.



Table 7

*Increased Instructional Strategies*

Instructional Strategy	Mean	Increase	Large Inc.	% Inc.	% Same
Writing assignments	3.87	52.63%	17.37%	70.00%	29.47%
Use of rubrics or scoring guides	3.79	48.42%	15.26%	63.68%	36.32%
Open response questions	3.71	42.86%	15.87%	58.73%	38.10%
Computers/educational software	3.71	44.51%	14.84%	59.34%	37.36%
Computers/internet/ research	3.70	44.51%	13.74%	58.24%	40.11%
Creative/critical thinking	3.68	44.92%	12.30%	57.22%	41.18%
Problem-solving activities	3.59	36.65%	13.61%	50.26%	45.03%
Facilitating/coaching	3.56	33.53%	11.38%	44.91%	54.49%
Calculators	3.55	29.06%	14.53%	43.59%	53.85%
Inquiry/Investigation	3.54	42.02%	7.45%	49.47%	47.34%
Visual aids (e.g. posters, graphs)	3.49	36.22%	6.49%	42.70%	56.76%
Supplementary books	3.46	41.30%	4.89%	46.20%	48.91%
Modeling	3.46	33.51%	6.81%	40.31%	58.64%
Use of response journals	3.46	35.48%	7.74%	43.23%	51.61%
Charts, webs, and/or outlines	3.46	37.10%	5.38%	42.47%	55.38%
Cooperative learning/group work	3.45	34.02%	8.25%	42.27%	52.06%
Interdisciplinary instruction	3.40	31.07%	7.34%	38.42%	55.93%
Collaborative/ team-teaching	3.39	31.06%	8.07%	39.13%	53.42%
Project-based assignments	3.37	36.13%	4.71%	40.84%	50.79%
Lab equipment	3.35	25.60%	5.60%	31.20%	67.20%
Audiovisual materials	3.34	31.02%	5.88%	36.90%	55.61%
Peer or cross-age tutoring	3.32	26.75%	5.10%	31.85%	63.69%
Newspaper/magazines	3.31	30.06%	5.20%	35.26%	56.65%
Use of manipulatives	3.30	28.14%	3.59%	31.74%	62.87%
Use of portfolios	3.30	25.32%	6.96%	32.28%	58.86%
Manipulatives	3.28	24.18%	3.92%	28.10%	67.97%
Discussion groups	3.27	26.98%	4.76%	31.75%	60.32%
Reference books	3.26	27.66%	3.19%	30.85%	60.64%
Maps/globes/atlas	3.23	19.84%	3.17%	23.02%	73.81%
Group projects	3.22	24.87%	4.76%	29.63%	58.20%
Primary source material	3.21	21.74%	1.09%	22.83%	74.46%
Lesson based on current events	3.19	28.24%	1.76%	30.00%	59.41%
Multiple-choice questions	3.15	22.34%	3.72%	26.06%	60.64%
Use of exhibitions	3.13	18.54%	1.32%	19.87%	71.52%

Also included on table 7 is the reported percentage of respondents reporting no change in use of instructional strategies and instructional materials. These percentages of no change in each of the survey items range from as low as 29.47% up to 71.52%.

Principals indicate observing an increase in the use of differentiated instruction techniques (schools A and Y), rubrics (schools A, C, Y), open-ended questions (schools A, C), writing prompts and/or journaling (schools A, C), reading strategies (schools C, Y, Z), chunking, and pre-assessment strategies (school Y). As to material usage, principals report an increase in the use of PSSA related items (schools A, Y, Z), writing prompts (A, Y), calculators, graphic organizers, and formula cards (C).

Of the 40 survey items addressing and instructional strategies and instructional materials, there were, in addition to those demonstrating an increase in use, a smaller number of items that indicated a decrease (mean <3.0) in use. The item that was designated as having the greatest decrease in use was lecturing. This item showed a total decrease in use of 28.13%. As indicated in Table 8, five other survey items, true-false questions (26.37%), textbooks (26.37%), worksheets (17.68%), textbook based assignments (18.44%), and role-playing (18.30%) demonstrated a decrease in use.

Reports from the principals indicated a decrease in the use of lecturing (schools A and Y), note taking (school Y), drill and practice (school C), and worksheets (school A). Another consequence reported by principals included the reduction of elective courses (school C) available to students.

Table 8

*Same or Decreased Instructional Strategies*

Instructional Strategy	Mean	% Decrease	% Large Decrease	Total Decrease	% Same
Role playing	2.99	15.03%	3.27%	18.30%	62.75%
Text-book assignments	2.98	16.76%	1.68%	18.44%	64.80%
Work sheets	2.95	14.58%	2.60%	17.19%	69.27%
Textbooks	2.87	16.02%	1.66%	17.68%	76.24%
True-false questions	2.76	21.43%	4.95%	26.37%	66.48%
Lecturing	2.74	26.04%	2.08%	28.13%	67.71%

Table 8 also shows the percentages of no change in use of instructional strategies and material as indicated by the respondents to the survey. The percentages range from a low of 62.75 to 67.71.

Instructional Strategies in High and Low performing Schools

Responses to Part I of the survey instrument were used in answering the second research question, how high-stakes testing results (PSSA Math and Reading) and the requirements of NCLB affected instructional strategies in high and low performing high schools? Overall, the teachers from both the high and low performing schools indicated that the change in use of instructional strategies and instructional materials to be similar. Table 9 shows the mean responses of the teachers from the high performing schools to each of the 51 survey items, as well as the mean responses to the same survey items from the low performing schools. The last column represents the difference between the high and low means.

Table 9

*High / Low mean differential*

Item	High	Low	Difference	Item	High	Low	Difference
1	3.769	3.988	-0.219	26	3.326	3.324	0.002
2	3.308	3.118	0.190	27	3.500	3.634	-0.134
3	2.980	2.975	0.005	28	2.909	2.817	0.092
4	3.379	3.151	0.228	29	3.194	3.329	-0.135
5	3.019	3.318	-0.299	30	3.406	3.530	-0.124
6	3.702	3.718	-0.016	31	3.178	3.253	-0.075
7	2.711	2.824	-0.113	32	3.312	3.313	-0.001
8	3.266	3.342	-0.076	33	3.481	3.169	0.312
9	3.621	3.435	0.186	34	3.420	3.268	0.152
10	3.712	3.448	0.264	35	3.350	3.754	-0.404
11	2.913	3.000	-0.087	36	3.775	3.625	0.150
12	3.269	3.104	0.165	37	3.772	3.617	0.155
13	3.394	3.345	0.049	38	3.289	3.271	0.018
14	3.765	3.576	0.189	39	3.217	3.246	-0.029
15	3.070	2.881	0.189	40	3.441	3.542	-0.101
16	3.433	3.488	-0.055	41	3.849	3.516	0.333
17	3.476	3.437	0.039	42	3.642	3.330	0.312
18	3.289	3.309	-0.020	43	3.632	3.747	-0.115
19	3.689	3.908	-0.219	44	3.991	4.044	-0.053
20	3.088	3.169	-0.081	45	4.123	4.088	0.035
21	3.237	3.588	-0.351	46	3.594	4.121	-0.527
22	2.743	2.747	-0.004	47	3.377	3.714	-0.337
23	3.481	3.437	0.044	48	3.547	3.667	-0.120
24	3.486	3.404	0.082	49	3.840	3.824	0.016
25	3.299	3.486	-0.187	50	3.179	3.396	-0.217
				51	3.868	4.156	-0.288

The high performing school mean and low performing school mean of the separate survey items, as illustrated in table 8, were very similar. Three of the items, however, indicated that teacher responses from the low performing schools demonstrated a greater increase in use of the instructional strategies and instructional materials than was demonstrated from the high performing schools. These included multiple-choice

questions (item #5), with a p-value equal to 0 .0022, interdisciplinary instruction (item #21), with a p-value equal to 0.0007, and calculators (item #35), with a p-value equal to 0.0030. Conversely, problem-solving activities (item #10) with a p-value equal to 0.0106 and audiovisual materials (item #33), with a p-value equal to 0.0021, demonstrated a significant increase in use by teachers in high performing schools when compared to low performing schools.

Table 10

*Significant difference of means 2-sample t-test*

Item #	p-value	Performance
Interactions with Colleagues	.0001	Low performance > High
Interdisciplinary instruction	.0007	Low performance > High
Audiovisual Materials	.0021	High performance > Low
Multiple-choice questions	.0022	Low performance > High
Calculators	.0030	Low performance > High
Interactions with principal(s)	.0036	Low performance > High
Personal desire to make changes	.0075	High performance > Low
Problem-solving activities	.0106	High performance > Low
Belief that changes will benefit students	.0171	High performance > Low

A chi-square analysis was performed between the items from the high performing school schools showing increased use and the survey items from the low performing schools showing increased use. The analysis, as shown in table 11, resulted in a p-value = 0.139, indicating no significance difference between the two sets of data.

Table 11

*High-Low Chi-square*

Item #	High Performing	Low Performing	Survey Item (cont.)	High Performing (cont.)	Low Performing (cont.)
1	0.101	0.113	27	0.028	0.031
2	1.232	1.376	28	0.563	0.629
3	0.087	0.097	29	2.419	2.702
4	0.264	0.295	30	0.421	0.470
5	1.817	2.030	31	0.451	0.504
6	0.066	0.074	32	0.044	0.049
7	0.003	0.003	33	3.207	3.582
8	0.217	0.242	34	0.434	0.485
9	1.016	1.135	35	2.325	2.598
10	0.978	1.093	36	0.658	0.736
11	0.364	0.407	37	0.565	0.631
12	0.355	0.396	38	0.123	0.138
13	0.090	0.100	39	0.032	0.036
14	1.042	1.164	40	0.033	0.036
15	0.298	0.332	41	0.968	1.081
16	0.065	0.073	42	0.705	0.788
17	0.040	0.044	43	0.180	0.201
18	0.028	0.032	44	0.000	0.000
19	0.366	0.408	45	0.081	0.090
20	0.340	0.379	46	1.056	1.180
21	2.474	2.764	47	1.226	1.369
22	0.144	0.160	48	0.011	0.012
23	0.051	0.057	49	0.009	0.010
24	0.340	0.380	50	0.760	0.849
25	0.545	0.608	51	0.092	0.103
26	0.051	0.057			

DF = 50      p = 0.139

The same analysis was performed between the high and low performing schools, but this included the data that indicated no change or a decrease of the 51 survey items (Table 12).

Table 12

*Chi-Square Test: High vs. Low, No change or Decrease*

Survey Item #	High Perf. Same & Decrease	High Perf. Same & Decrease	Low Perf. Same & Decrease	Low Perf. Same & Decrease	Total
	Observed	Expected	Observed	Expected	Observed
1	35	31.01	21	24.99	56
2	67	73.10	65	58.90	132
3	82	81.96	66	66.04	144
4	67	70.88	61	57.12	128
5	82	75.86	55	61.14	137
6	42	42.64	35	34.36	77
7	89	93.03	79	74.97	168
8	67	62.57	46	50.43	113
9	45	52.05	49	41.95	94
10	45	52.61	50	42.39	95
11	91	91.37	74	73.63	165
12	64	66.45	56	53.55	120
13	59	62.02	53	49.98	112
14	36	43.75	43	35.25	79
15	68	68.11	55	54.89	123
16	62	58.70	44	47.30	106
17	47	48.73	41	39.27	88
18	63	59.25	44	47.75	107
19	43	38.21	26	30.79	69
20	67	67.00	54	54.00	121
21	69	59.81	39	48.19	108
22	99	101.34	84	81.66	183
23	61	63.13	53	50.87	114
24	57	62.02	55	49.98	112
25	56	53.71	41	43.29	97
26	61	59.25	46	47.75	107
27	56	50.39	35	40.61	91
28	90	93.58	79	75.42	169
29	79	70.88	49	57.12	128
30	59	54.82	40	44.18	99



Table 12 (cont.)

*Chi-Square Test: High vs. Low, No change or Decrease*

Survey Item #	High Perf. Same & Decrease	High Perf. Same & Decrease	Low Perf. Same & Decrease	Low Perf. Same & Decrease	Total
	Observed	Expected	Observed	Expected	Observed
31	80	77.53	60	62.47	140
32	61	61.47	50	49.53	111
33	56	65.34	62	52.66	118
34	43	46.52	41	37.48	84
35	39	35.44	25	28.56	64
36	38	40.98	36	33.02	74
37	38	41.53	37	33.47	75
38	61	60.91	49	49.09	110
39	54	54.27	44	43.73	98
40	60	58.14	45	46.86	105
41	21	29.35	32	23.65	53
42	35	42.09	41	33.91	76
43	43	39.87	29	32.13	72
44	18	17.17	13	13.83	31
45	10	12.18	12	9.82	22
46	34	24.37	10	19.63	44
47	50	41.53	25	33.47	75
48	36	35.99	29	29.01	65
49	26	26.03	21	20.97	47
50	64	59.81	44	48.19	108
51	27	23.26	15	18.74	42

P = 0.303

The analysis was set up between the high performing schools no change or decrease of use and the low performing schools no change or decrease of use. The analysis, as demonstrated in table 12, resulted in a p-value = .303, indicating no significant difference in the data sets.

A third procedure was completed to ascertain differences or similarities between the high and low performing schools. A regression analysis was completed comparing the combined mean of the three high performing schools to the combined mean of the three low performing schools. The resulting coefficient of determination (see table 13) equaled 73.9% ( $r^2 = 73.9\%$ ).

Table 13

*High-Low Curriculum*

The regression equation is Low Mean = 0.392 + 0.883 High Mean

Predictor	Coef	SE Coef	T	P
Constant	0.3922	0.2589	1.52	0.136
High Mean	0.88305	0.07504	11.77	0.000

S = 0.1685    R-Sq = 73.9%    R-sq (adj) = 73.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	3.9305	3.9305	138.49	0.000
Residual	49	1.3906	0.0284		
Factor					
Total	50	5.3212			

Unusual Observations

Survey Item	High Mean	Low Mean	Fit	SE Fit	Residual	St Residual
True-false questions	2.72	2.82	2.7941	0.0587	0.0259	0.16 X
Lecturing	2.72	2.74	2.7941	0.0587	-0.0541	-0.34 X
Interest in helping students attain PSSA scores allowing them to graduate	4.16	4.07	4.0657	0.0592	0.0043	0.03 X
Interest in avoiding sanctions	3.65	4.12	3.6154	0.0285	0.5046	3.04 R

R denotes an observation with a large standardized residual  
 X denotes an observation whose X value gives it large influence

This item-by-item comparison between the high and low performing schools demonstrates that the teacher responses indicate that both groups have changed in a fairly similar manner.

Table 14 illustrates the mean responses to each of the 51 survey items, of the English and math teachers from all six of the participating high schools.

Table 14

*Math and English Teachers: Mean Responses*

Item #	English Mean	Math Mean	Difference	Item # (cont.)	English Mean (cont.)	Math Mean (cont.)	Difference (cont.)
1	3.947	3.794	0.153	27	3.677	3.438	0.240
2	2.842	3.171	-0.329	28	2.778	2.971	-0.194
3	2.784	2.943	-0.159	29	2.946	3.235	-0.289
4	3.079	3.182	-0.103	30	3.395	3.629	-0.234
5	3.105	3.500	-0.395	31	3.184	3.059	0.125
6	3.789	3.971	-0.181	32	3.027	3.115	-0.088
7	2.459	3.030	-0.571	33	2.946	3.344	-0.398
8	2.913	3.314	-0.401	34	3.200	3.321	-0.121
9	3.583	3.514	0.069	35	3.000	3.912	-0.912
10	3.405	3.771	-0.366	36	3.457	3.657	-0.200
11	2.737	3.200	-0.463	37	3.553	3.529	0.023
12	2.813	3.160	-0.348	38	2.947	3.314	-0.367
13	3.211	3.206	0.005	39	2.944	3.087	-0.143
14	3.632	3.788	-0.156	40	3.368	3.471	-0.102
15	2.618	2.889	-0.271	41	3.795	3.800	-0.005
16	3.405	3.314	0.091	42	3.436	3.600	-0.164
17	3.622	3.292	0.330	43	3.897	3.800	0.097
18	3.270	3.357	-0.087	44	4.231	4.086	0.145
19	3.921	3.771	0.150	45	4.333	4.057	0.276
20	3.000	3.042	-0.042	46	4.077	3.771	0.305
21	3.000	3.313	-0.313	47	3.513	3.571	-0.059
22	2.595	2.743	-0.148	48	3.590	3.771	-0.182
23	3.605	3.400	0.205	49	3.821	3.914	-0.094
24	3.105	3.543	-0.438	50	3.103	3.229	-0.126
25	3.071	3.467	-0.395	51	4.051	4.000	0.051
26	3.080	3.419	-0.339				

Table 15

*Significant Difference 2 sample T test*

Survey Item #	p value	
Multiple-choice questions	0.0246	Math > Eng
True-false questions	0.0004	Math > Eng
Use of Manipulatives	0.0053	Math > Eng
Problem-solving activities	0.0278	Math > Eng
Work sheets	0.0014	Math > Eng
Cooperative learning	0.0020	Math > Eng
Collaborative/team teaching	0.0259	Math > Eng
Audiovisual materials	0.0094	Math > Eng
Manipulatives	0.0062	Math > Eng
Use of response journals	0.0257	Eng > Math
Interest in avoiding sanctions	0.1019	Eng > Math

Also included in the table is the difference in mean responses between the participating English and Math teachers. These calculations of mean responses of the English and math teachers were used to complete a significant difference of means, 2-sample T-test (Table 15). The completed test indicated 10 survey items that demonstrated a significant difference between the two groups. Nine of the 10 items indicated a greater increase in the use of instructional strategies and materials by the Math teachers. These included multiple choice questions (p-value = 0.0246), true-false questions (p-value = 0.0004), use of manipulatives (p-value = 0.0053), problem-solving activities (p-value = 0.0278), worksheets (p-value = 0.0014), cooperative learning/group work (p-value = 0.002), collaborative/team teaching (p-value = 0.0259), audiovisual materials (p-value = 0.0094), and manipulatives (p-value = 0.0062). Only one item, use of response journals (p-value = 0.0257) demonstrated a greater increase in use by the English teachers as opposed to the math teachers.

## Curriculum Changes

The principals of the participating schools provided the information necessary to answer the third research question, how do high-stakes testing results affect curriculum in the high and low performing schools?

The principals from the schools involved in the study offered their opinions concerning the yearly administration of the PSSA reading and math exams. Viewpoints of the principals from the high performing schools differed from that of the principals from the low performing schools. The principals that reported the least amount of impact, schools A and B, both from high performing schools related a small amount of change to the curriculum in their schools to the administration of the testing and then reporting of the PSSA student and school scores. Many of the curricular changes in these two schools were made a few years earlier during restructuring of the master schedule and the implementation of the block schedule. The only curricular changes they discussed were those involving the lowest achieving students in their schools. Changes included the addition of remediation courses in reading and math, specifically intended to aid those students scoring below the proficient level on the assessment. The only other changes that were discussed by all three of the principals from the high performing schools included alignment of each curriculum to the mandated state standards. The principals from the low performing schools also discussed the alignment of curriculum to the state standards. But they also talked about the incorporation of the state supplied anchors into the regular delivery of the content as well as the ongoing practice of the testing strategies. The anchors are content and test specific samples of the types of

questions that will appear on the next test. They also provide a guide to the content areas that will be tested. These principals indicated that testing has increased well beyond that of the PSSA exams. These schools have implemented the use of two other diagnostic tests, administered to the students each of the four years in the high school. Along with the additional testing, students are provided with additional test preparation lessons, as each test is different in design. This time on testing and preparation for testing has a huge impact, reducing instructional time.

Principals from a number of the schools, both high and low performing, briefly discussed the process of comparing the achievement and quality of schools to other schools based on a single test. They all talked about a need to change the accountability system to focus on individual student growth from year to year.

### Principal Perceptions

Interviews provided information relative to how principals in high and low performing high schools view the PSSA standardized testing and accountability program. The three principals from the high performing schools expressed their satisfaction in having positive school scores. Each of their schools has been recognized as having met the Pennsylvania Department of Education standard of annual yearly progress. Additionally, these schools were three of the five top performing schools in the four county region of northwest Pennsylvania. Descriptors of the testing program, as provided by the principals from high performing schools included inconvenient, excessive, and disruptive. They described the testing of individual students followed by the reporting of school progress as being poorly designed. They felt that the comparison from school to

school has been unfair, as has been the comparison of the achievement scores of one group of students a second group.

The principals from the lower performing schools offered similar responses in terms of the test being an unfair comparison of student and school scores. They also talked about the need to change the system of year-to-year comparison of student progress that does not include progress of individual students over multiple years. However, these principals also indicated that the testing and reporting program has provided an opportunity for their schools to improve the education in their school. With low performance scores, the schools have been able to offer additional support for the students and teachers in the classrooms by increasing the professional and support staff personnel.

### Influence on Changes

Responses to Part Two of the survey instrument provided information to answer the last research question investigating what factors have influenced changes (in use of instructional strategies and materials) in high and low performing high schools? Teachers were asked to indicate agreement or disagreement to 11 survey statements concerning the impetus for change in the classroom. Combined responses from teachers in all six high schools demonstrated that 88.83% of the teachers (mean = 4.11) were motivated by an interest in helping students to succeed, followed by 84.77% (mean = 4.02) indicating an interest in helping the school to improve PSSA scores, 78.57% (mean = 4.00) reporting alignment of curriculum to the state standards as a factor, and 77.66% (mean = 3.84) indicated that interest in avoiding sanctions was a motivation for change in



the classroom. Conversely, interactions with parents, was assigned the lowest level of agreement with only 45.18% (mean = 3.28). Complete mean and percentage data for Part II can be found in Table 16.

Table 16

*Influence Factors*

Instructional Strategy	Mean	Total % Agree	Total % Disagree
Personal desire to make changes	3.69	72.59%	14.21%
Belief that such changes will benefit students	3.50	61.42%	19.80%
Changes in the types of assessment used for school accountability	3.69	63.45%	6.09%
Interest in helping my school improve PSSA scores	4.02	84.77%	6.09%
Interest in helping my students attain PSSA scores that will allow them to graduate	4.11	88.83%	3.55%
Interest in avoiding sanctions at my school	3.84	77.66%	12.69%
Interactions with school principal(s)	3.53	61.42%	14.21%
Interactions with colleagues	3.60	66.84%	12.76%
Staff development in which I have participated	3.83	76.65%	8.12%
Interactions with parents	3.28	45.18%	21.83%
Curriculum was aligned to coordinate with state standards	4.00	78.57%	5.10%

Principals from the three high performing schools report the amount of pressure to change placed on the school to be little to none. They indicate that teachers are provided with PSSA testing anchors that act as a guide to the concepts that will be assessed on the state exam. Students in school A and school B are held to school district requirements that place students performing below the proficient level on the state exams in remedial courses.

Conversely, the principals from the low performing schools report pressure from the school administration on the teachers (school Y) to use the state provided assessment anchors, reminders from the school board and/or district administration (school X) concerning sanctions for poor performance on the exams.

Part II of the survey was included in the regression analysis (low mean vs. high mean) illustrated in Figure 1, earlier in this chapter. The purpose was to ascertain differences or similarities between the high and low performing schools. The regression analysis compared the combined mean of the three high performing schools to the combined mean of the three low performing schools. The resulting coefficient of determination (see Table 13) equaled 73.9% ( $r^2 = 73.9\%$ ). This item-by-item comparison between the high and low performing schools indicated that both teacher groups have changed the use of instructional strategies and materials in a fairly similar manner. However, interest in avoiding sanctions was identified as significant as an outlier, not fitting in the same pattern as the other items on the survey. The mean response from the low performing schools (4.12) was significantly higher than that of the high performing schools (3.65), and was three standard deviations away from the regression line.

#### Instructional Practices and Years of Experience

To examine how the changes in use of instructional strategies and materials were affected by the years of experience as reported by the participants, a chi-square test was completed comparing Part I (survey items 1-40), to Part III, item 53, teaching experience (see Table 17).

Table 17

*Years Experience vs. Change*

Years of Experience	High Perf. No change or Decrease Observed Counts	High Perf. No change or Decrease Expected Counts	Low Perf. No Change or Decrease Observed Counts	Low Perf. No Change or Decrease Expected Counts	Total Observed Counts
3 or less	326	281.54	317	361.46	643
4-12	1022	972.02	1198	1247.98	2220
13-19	523	643.20	946	825.80	1469
20-27	594	583.65	739	749.35	1333
28 or more	596	580.59	730	745.41	6991

p-value = 0.0000

Table 18

*Years Experience vs. Instructional Strategies*

Years Exp.	Sum	n	Mean 1-40
1	2255	643	3.506998
2	7346	2220	3.309009
3	4819	1469	3.280463
4	4432	1333	3.324831
5	4547	1326	3.429110

The results of the chi-square test indicated that change of use was dependent on the years of experience (p-value=0.0000). As displayed in Tables 17 and 18, two groups of teacher participants indicated more changes in use of instructional strategies and materials than the other teacher participants. Teacher participants reporting one to three years of experience reported the highest level of change (mean=3.50), followed by the most experienced participating teachers with 28 or more years (mean=3.42) in the classroom.

The remaining teachers, reporting 4 to 12 (mean=3.30), 13-19 (mean=3.28), and 20-27 (mean=3.42) years of experience, reported the least amount of change in the classroom.

### Instructional Practices and Level of Education

To examine how the changes in use of instructional strategies and materials were impacted by the teachers' level of education, as reported by the participants, responses to survey item 54, education, was compared to survey items 1-40. The reported levels of education were divided into two groups, bachelor degree and master's degree. The one teacher that had completed the doctorate was not included in either group. The response from the teachers having completed a bachelor's degree equaled a mean of 2.93, while responses from those teachers that have completed a master's degree equaled a mean of 3.01. Simply put, the teachers holding a masters degree indicated more change in the use of instructional strategies and materials than the teachers holding a bachelor degree. Those with a bachelor degree, with a mean below 3.0, actually indicated a small decrease in the use of instructional strategies and materials.

Using the same data set, a chi-square test was performed comparing Part I (survey items 1-40), to Part III (survey item 54). Even though the difference between the two means was relatively small (2.932 as compared to 3.011), the resulting p-value was equal to 0 .005 indicating the difference in means was significant (see Table 19).

Table 19

*Educational Level vs. Part I, Part II*

Education Level	Sum 1-40	n	Mean	Sum 41-51	n	Mean
Bachelors	12077	4120	2.931311	4160	1133	3.671668138
Masters/Doctorate	11322	3760	3.01117	3926	1034	3.796905222
*Note: only 1 doctorate						
Standard Deviation						
Bachelors			1.269046			0.926571126
Masters/Doctorate			1.255319			0.879667112
Significance Test						
Questions 1-40	0.005					
Questions 41-51	0.001					

Influence

To examine how the influence for change was affected by the teachers' reported level of education, responses to survey item 54, level of education, was compared to survey items 41-51. Once again the levels of education were divided into two groups, bachelor degree and master degree. The mean responses of the teachers that held master degrees equaled 3.79, while the mean responses from the teachers with bachelor degrees equaled 3.67. A chi-square test (see Table 19) was performed comparing Part I (survey items 41-51) to Part III (survey item 54). The resulting p-value was equal to 0.001 indicating significance between the factors influencing change and the level of education achieved by the teachers in the study.

To examine how the reasons for change were affected by the years of experience, as reported by the participants, survey items 41-51 were compared to the years of

experience. The teachers that have been in the classroom for 13 to 19 years indicated agreement, at a rate of 64.5%, with the statements describing influences causing change in instructional strategies. Those working 4-12 years reported 67.6% agreement, three or less years of teaching experience indicated 73.6% agreement, 20-27 years experience indicated 73.7% agreement, and the most experienced teachers, in the classroom for 28 or more years indicated the highest level of agreement at 77.9%. See Table 20 for complete details comparing years of experience to the factors influencing change.

Table 20

*Years of experience vs. influence for change*

Years Exp.	SD	D	U	A	SA	% Agree	% Disagree
1	0	21	40	114	56	0.736	0.091
2	23	63	138	355	113	0.676	0.124
3	9	57	90	226	58	0.645	0.150
4	6	31	67	220	71	0.737	0.094
5	0	35	55	271	46	0.779	0.086

A chi square test was performed comparing the total responses indicating disagreement, to those representing agreement to items 41 through 51 from the survey. Results, as seen in Table 21, include a p-value of 0.003, indicating that the years of experience reported by the teachers do affect how the survey items were answered. Responses to the survey items are dependent on the years of experience. It can also be noted that those reporting 13-19 years of experience indicated the highest level of disagreement.

Table 21

Years Experience compared to influence for change

Years Experience.	Total Disagree	Total Agree
<3	21	170
4-12	86	468
13-19	66	284
20-27	37	291
28+	35	317

Chi-Square Test

p =0.003

Largest contribution to test statistic is number of exp. 3 teachers who disagree (higher than normal).

## CHAPTER V: DISCUSSION

### Introduction

This chapter presents a summary of the study that investigated the resulting effects on classroom instruction as a consequence of the administration of the eleventh grade, PSSA reading and math exams. This chapter will review the research problem and the methodology used in the study. Important conclusions drawn from the data presented in Chapter IV will also be presented, along with a discussion of the implications for action and recommendations for future research.

### Summary of the Study

The purpose of this study was to investigate the role of the PSSA on the beliefs of teachers and principals, and the impact of the PSSA on the instructional program in their schools. More specifically, this study examined the instructional consequences realized between lower performing high schools and high performing high schools as reported by the Pennsylvania Department of Education.

In an attempt to hold students, teachers, and schools accountable, a large majority of schools, school districts, and states have implemented the use of norm-referenced standardized tests to measure student achievement. Forty-nine (49) states have allocated funding for the development of testing instruments to assess effectiveness of schools (Moon, Brighton, & Callahan, 2002).

Kohn (1999) suggests that the emphasis placed on test scores will eventually result in increased importance on classroom strategies that focus on the development of



basic skills and drill and practice. He also implied that high-stakes testing has resulted in a drastic change in the way instruction is delivered, and few if any of the changes have been positive. Schools and communities across the country have made it acceptable to provide drills and practice for a high-stakes test, even if the students are learning very little (Kohn, 2000).

The literature suggests that (a) classroom instruction has taken on the same format as the test, (b) tests and quizzes have taken on the form of multiple-choice items and the activities associated with learning have also become multiple-choice, (c) essay tests have been abandoned because they provide no advantage to the student taking the standardized test, (d) the exam has determined what was to be included and excluded from the curriculum, (e) the teaching and learning process has been narrowed to the content of the test, (f) attention to specific skills necessary to score higher on a test have become more important than the more difficult skills such as critical thinking and problem solving (Green, 1991; Kohn, 1999). Kohn (1999) also implied that these tests could not take into account the process used by the test taker to arrive at any given answer, therefore ignoring the question of understanding.

Other possible consequences of high-stakes testing may include the narrowing of instruction; weakening of pedagogy; declining innovations; a shift from student-centered to teacher-centered classrooms (Saltman & Gabbard, 2003); unhealthy, unethical competition among teachers; coaching students during exams; changing answers on the exam (Wilde, 2004; Kohn, 2000); alienation of student groups; grade-level retention; and teaching to the test (Wilde, 2000). In addition, teachers may begin to dislike having the

lower level student in their classes because they might reduce the perceived success of the teacher (Kohn, 2000).

Alternatively, standardized testing can produce useful information when used as intended. Standardized, norm-referenced tests were designed to provide educators with information about how well the students achieve in academic areas, as compared to similar groups of students. However, with the current use of standardized, norm-referenced tests, individuals and schools are being held accountable for factors beyond their power to control, such as the number of students identified as special needs, level of impairment and need, or economically disadvantaged. Low scores have often been connected to the socioeconomic surroundings of the student, which have included the school resources as well as the affluence of the school community.

The Department of Education has implemented the Pennsylvania System of School Assessment (PSSA). Students are required to score at a state-mandated level of proficiency in reading, and mathematics before becoming eligible to receive a high school diploma. And more recently, the federal government, through NCLB legislation, has required schools, by the year 2014, to have all students proficient in reading, and mathematics. With this mandate also comes an increased amount of testing for students (National Association of Elementary School Principals & National Association of Secondary School Principals, 2003).

This study collected data on (a) teachers' perceptions of the PSSA standardized testing and accountability system; (b) how the program affected instructional practices, and to what extent teachers changed their instructional practices as a result of PSSA score reporting and expected improvement; and (c) what factors, if any, may have influenced

these changes. The researcher intended to investigate (a) how the PSSA math and reading assessments, since the reporting of scores and institution of NCLB, impact instruction and instructional strategies? (b) how high-stakes testing results (PSSA Math and Reading) and the requirements of NCLB affect instructional strategies in high and low performing high schools? (c) how high-stakes testing results (PSSA Math and Reading) affect curriculum in high and low performing high schools? (d) how principals in high and low performing high schools view the PSSA standardized testing and accountability program? and (e) what factors influenced these changes in high and low performing high schools?

### Methodology

The study included certified teachers and principals from six secondary schools (grades 9-12) located in the northwestern region of Pennsylvania. Schools were selected based on the test results from the 2004-2005 school year administration of the PSSA 11th grade reading and math exams as reported by the Pennsylvania Department of Education. Three of the six high schools included in the study were high performing schools, or schools realizing scores above the 70 percent mark of students at the proficient or advanced level in both math and reading. The remaining three schools were included based upon scores of 35 percent or higher number of students scoring below proficient on the latest administration of the PSSA reading and math exams. Schools were purposefully selected based on scores, location, and availability to the researcher.

The researcher was permitted to seek volunteers from each of the six high schools, in each of the specialized areas of English, Math, Science, Social Studies, Art,

Music, Physical Education, Business, Foreign Language and Industrial Arts. Volunteers were encouraged to complete the survey and return completed forms to the researcher. Teacher participants included 197 of 249 (79.11%) potential candidates, 91 from low performing schools (X, Y, Z), and 106 from the high performing schools (A, B, C). The response rate from the high performing schools (106 of 126 or 84.13%) was approximately 10% higher than from the low performing schools (91 of 123 or 73.98%), even though the staffs from the two sets of schools were somewhat similar in size (difference of three teachers). The differences between the low and high response rates are difficult to explain within the scope of this study. However, some of the curricular areas were represented with very different rates of participation between the low and the high performing schools. For example, 13 special education teachers from the low performing school submitted completed surveys where there were only six special education teachers representing the high performing schools. One would expect that the lower performing school might have a larger special needs population, or that the school deals with a higher percentage of disadvantaged students. Another noticeable difference was that the lower performing schools had five less teachers submit completed surveys from each of four areas of study, English, math, science and physical education. One might ascertain that the lower performing schools have a greater need to staff for special populations. Another possibility is that the school community has placed a lesser priority on either maintaining appropriate class size or the number of required courses needed for graduation in their respective school. Appropriate class size varies greatly from school district to school district as well as from building to building within a given district. A district that places a higher priority on smaller classes infers a priority on the delivery of

instruction, understanding of content, assessment, and achievement of the individual student. An increased credit requirement in the core curricular areas also infers a high priority to not only a solid foundation of content, but also a deeper understanding the leads to practical application of the content. Information detailing the specific number of required courses or credits from the curricula areas in English, math, and science might demonstrate the priority placed on math by the high performing schools. It is possible that the low performing schools have not increased the math or science requirement beyond the state minimum of three credits in four years of secondary school. And finally, it may be that teachers from the low performing school were either busy with students and could not attend the faculty meeting or were not interested in attending.

The 54-item, three-part survey asked the respondent to indicate to what extent he or she may have changed or altered the use of 40 instructional strategies and techniques used in their classroom. Section one asked the respondents to indicate a degree of change in usage for each item. Likert-style responses asked the respondent to indicate large increase in use (LD), decrease in use (D), same or steady use (S), increase in use (I), or large increase in use (LI). Respondents were also able to indicate that the item did not apply to their curricular area (NA).

Part II of the survey asked teachers to respond to 10 items indicating what factors influenced changes in educational practices. This portion of the survey required Likert-style responses ranging from strongly disagree to strongly agree. In section three of the survey, respondents were asked to supply demographic information describing sex, years of teaching experience, level of education or degree earned, and area of certification (Vogler, 2000).

The data was analyzed using Excel and Minitab statistical software. Responses to the survey questions in section one and two were analyzed using descriptive explanatory techniques reporting the mean and percentages of the responses (Fink, 1995).

## Changes to Instructional Practices (RQ-1)

Contrary to several theories presented in the literature (e.g. Green, 1991; Moore, 1994; Hoffman, 2001; Moon, 2004; Kohn, 1999) concerning the effects of testing on instruction, the results of this study indicate that participating teachers, from both low and high performing schools, believe that the PSSA math and reading assessments has led to change in their instructional practices since the implementation of NCLB and the reporting of PSSA scores. Classroom instruction has changed, but in a positive direction. Some of the claims against testing suggest that teachers have abandoned the use of essays, that classroom instruction has taken on the same format as the tests, that the teaching and learning process has been narrowed to the content of the test, that attention to specific skills necessary to score higher on a test have become more important than the more difficult skills such as critical thinking and problem solving (Green, 1991; Moore, 1994; Kohn, 1999; Hoffman, 2001). To the contrary, teachers report an increased use in critical thinking and problem solving strategies.

The teachers indicated increasing the use of 36 of the 40 instructional strategies identified on the survey. The items showing the largest percent of increase included, writing assignments (70%), rubrics and scoring guides (63.68%), and open ended questions (58.73%), all being necessary in the assessment of higher order skills (Nitko, 2004). The results of this study support a research by Vogler (2002) that presents the opposing view of many of the existing theories presented in the current literature. Volger (2002) identified these same items as instructional strategies showing the largest increase in use and advocated by the Massachusetts Department of Education.

Teachers have also decreased the use of six of the forty survey items. Reported decreases included lecturing (28.13%), true-false questioning (26.37%), textbook based assignments (18.44%), role-playing (18.30%), textbooks (17.68%), and worksheets (17.19%). Of these items, Vogler (2002) identified the same three items with the greatest amount of decrease of use in the Massachusetts study. Vogler also reports that these items are not advocated by the Massachusetts Department of Education (2002) and that these items fail to develop high order thinking skills. These results indicate a change in use of instructional strategies since the implementation of NCLB and reporting of the PSSA scores. The results suggest that instructional strategies have changed to include more appropriate pedagogical practices since the implementation of NCLB and the reporting of PSSA scores, contradicting statements made by Kohn (1999) that predict movement away from high order skills and toward the development of basic skills, and that teachers will present test taking skills rather than the approved curricular content, and Green (1991) describing the diversion away from critical thinking and problem solving skills to allow time for the teaching of test taking strategies. However, Kohn (2000), and Abrams et al. (2003) are supported by the results of this study in predicting the increase in the use of multiple-choice tests and quizzes as a practice for the administration of the PSSA exams.

#### Effect of Testing on Instructional Strategies (RQ-2)

To address the second research question, high-stakes testing results (PSSA Math and Reading) and the requirements of NCLB do affect instructional strategies in high and low performing high schools very similarly. In analyzing the results of the 2-sample T-



test comparing the mean responses from high performing schools to low performing schools, it appears that 35 of the 40 items were very similar, showing no significant difference in the responses. Five items however did show a difference worth noting, with the greater difference being demonstrated in the lower performing schools demonstrating a greater increase in use. This does not indicate a decline in use by the high performing schools. Both the high and low performing schools indicated increases, but the increase in the lower performing schools was demonstrated at greater degree. These included multiple-choice questions (p-value = 0.0022), interdisciplinary instruction (p-value = 0 .0007), and calculators (p-value = 0 .0030). It may be that the low performing schools perceived a need to have students practice testing in the same format as the state assessment. Unfortunately, a large percentage of the PSSA math and reading exams require students to provide answers to multiple-choice questions. These tests, usually used to measure basic comprehension and application only require that the student identify the requested information through matching with pre-existing memory (Svinicki & Koch, 1984). Multiple-choice tests are not designed to "...assess higher order thinking, problem solving abilities, creativity, or initiative" (Davey & Neil, 1991, p. 3). Baker and Hoffman (as cited by Haladyna, 1992), agree that multiple choice testing emphasizes such lower order learning skills as recall of facts. The low performing schools would also recognize the ability of interdisciplinary instruction to assist in the development of high order thinking skills, and at the same time have a larger capacity for growth. The high performing schools have already been practicing and improving on the use of interdisciplinary instruction. Conversely, problem-solving activities (p-value = 0 .0106) and audiovisual materials (p-value = 0.0021), demonstrated a significant increase

of use in the high performing schools as compared to the low performing schools. In regard to the increase of use in problem-solving activities, the high performing schools had no need to spend large amounts of time to reinforce basic skills in reading and math. It is possible that these schools have adjusted curriculum to require students to demonstrate mastery of the basic skills prior to the state testing. These high performing schools were instead able to spend that instructional time developing problem-solving activities.

### Effect on Curriculum (RQ-3)

The principals from the schools involved in the study offered their opinions concerning the yearly administration of the PSSA reading and math exams. As might be expected, the viewpoints of the principals from the high performing schools differed from that of the principals from the low performing schools. The principals from high performing schools reported the least amount of impact from the PSSA. Curricular changes in the high performing schools were minimal, but most purposely targeted to the lowest achieving students in their schools. Besides the alignment of the curriculum to the state standards, changes included the addition of reading and math remediation courses intended to aid students scoring below the proficient level on the assessment. The principals from the low performing schools also discussed the alignment of curriculum, but they also talked about the incorporation of the state supplied assessment anchors into the curriculum.

These discussions with principals indicate that the resulting consequences of the PSSA may be different for the high and low performing schools. The curriculum in the

high performing schools has been changed very little as a result of the state-testing program. The students and schools have scored relatively high as compared to the many schools in northwestern Pennsylvania. These schools have been able to maintain the local developed and implemented curriculum while adjusting course offerings to include courses to improve achievement levels of the lower performing students in the tested grade levels. The lower performing schools on the other hand are in a different situation. These schools, as are many others in the state, have been identified as schools in warning, or have been placed on a watch list by the state department of education. These schools have been and continue to be required to make significant gains in achievement in order to reach state determined minimum school scores. They have also been asked to submit an annual plan of action indicating changes that will be implemented as they address deficiencies in the schools. Curriculum in these schools, X, Y, and Z, as reported by the principals, has changed to a greater degree. And because of the large numbers of students performing below the proficient levels on the exams, many of the reading and math course offerings are aligned to the content included on the state test.

#### Principal's View of PSSA (RQ-4)

How do principals in high and low performing high schools view the PSSA standardized testing and accountability program? Interviews provided information relative to how principals in high and low performing high schools view the PSSA standardized testing and accountability program. As one would expect, the three principals from the high performing schools expressed satisfaction in having positive school scores, and being recognized as three of only five top performing schools in the

four county region. These principals described the PSSA testing program as inconvenient, excessive, and disruptive to the learning environment, poorly designed, and unfair in the comparison of schools. The principals from the lower performing schools offered similar responses in terms of the test being an unfair comparison of student and school scores, but also talked about the need to change the testing and reporting system to one that includes a year-to-year comparison of individual student progress over multiple years. However, these principals, including one from an urban setting, also indicated that the testing and reporting program did provide opportunity for their schools to improve the delivery of education in their school. The mandate of NCLB and adequate yearly progress provided the motivation for the entire school community to improve the delivery of education and the achievement level of the students. The fear of failure to meet annual yearly progress, coupled with the eventual sanctions that could be imposed by the department of education, was the much-needed incentive to change instructional practices.

### **Influence for Change (RQ-5)**

Part two of the survey instrument provided information to answer the last research question investigating what factors have influenced changes (in use of instructional strategies and materials) in high and low performing high schools? Teachers were asked to indicate agreement or disagreement to 11 survey statements concerning the impetus for change in the classroom.

As also reported in the study by Reed et al. (2001), principals from the three high performing schools report the amount of pressure to change placed on the school to be little to none. They indicate that teachers are provided with PSSA testing anchors that act

as a guide to the concepts that will be assessed on the state exam. Students in school A and school B are held to school district requirements that place students performing below the proficient level on the state exams in remedial courses.

Reed et al (2001) is also supported by the results of this study that more pressure is placed on the teachers by the principals and administration from the low performing schools (school Y) to use the state provided assessment anchors, reminders from the school board and/or district administration (school X) concerning sanctions for poor performance on the exams.

Part II of the survey was included in the regression analysis (low mean vs. high mean) illustrated in Figure 1, earlier in this chapter. The purpose was to ascertain differences or similarities between the high and low performing schools. The regression analysis compared the combined mean of the three high performing schools to the combined mean of the three low performing schools. The resulting coefficient of determination (see Table 13) equaled 73.9% ( $r^2 = 73.9\%$ ). This item-by-item comparison between the high and low performing schools indicated that both teacher groups have changed the use of instructional strategies and materials in a fairly similar manner. However, interest in avoiding sanctions was identified as significant as an outlier, not fitting in the same pattern as the other items on the survey. The mean response from the low performing schools (4.12) was significantly higher than that of the high performing schools (3.65), and was three standard deviations away from the regression line.

The motivation for the teachers involved in this study to change their instructional practices in the classroom appears to come from a genuine interest in wanting the students to succeed, the desire to help the school improve the PSSA scores, and to avoid

state imposed sanctions for poor performance scores. This would fit with what one would expect from teachers, regardless of the performance level of the school. Students, parents, and the general public have a perception that teachers are good people, interested in the success of the students and school. This is supported by some of the other survey items that suggest a belief in helping students to succeed. The teachers indicated that they believed that by changing the way they delivered the education, students would achieve at a higher level, and advance to graduation. The teachers also indicated that the desire to change was motivated by the education received through staff development programs. They most likely recognized the positive effect of the programs and alignment of the curriculum to the state standards.

In the comparison of the two groups of schools, low performing schools indicated a high level of pressure being exerted by the school administration. As might be expected, the teachers from the high performing schools realize little to no pressure at all. Again this would not be a surprise, as the lower the school score, and the longer those scores remain low, the higher the possibility of the schools being required to accept sanctions from the state department of education.

When the teachers' level of education was compared to influence to change, calculations show a strong significance between the factors. Teachers with the higher level of education, master degree and above, were more likely to institute change in the use of instructional strategies than those teachers with less education. This could be explained in two ways. The first is that the teachers with less experience, coming from teacher preparation programs are already using the strategies that relate to higher order thinking skills. Therefore there would be little need to change their practices. On the

other hand, their limited education and experience may not have been enough to realize that the students may need different strategies. The second scenario presents the possibility that the more experienced teachers have more to change. The large majority of these teachers also have more years of experience in the classroom. This experience, combined with the new testing reporting requirements, may suggest that the strategies that they have been using needs adjustment. In either scenario, the more experienced and more educated the teacher, the more likely that the school will realize a change in the delivery of instruction.

### Limitations

This study was limited to six high schools in northwestern Pennsylvania. The schools were selected based on the 2004 PSSA reading and math scores as distributed by the Pennsylvania Department of Education. Other factors that were considered when selecting these schools included location in respect to that of the researcher, as well as availability to the faculties of the selected schools. The researcher did not experience any difficulties beyond those encountered when attempting to gain access to high and low performing schools in one urban school district.

One area, in which this study was limited, was in the collection of socio-economic data of the participating school communities. Data to be collected in future research would include the number of students receiving free or reduced lunch and/or breakfast, the size of the special education identified population, as well as the number of students enrolled in the English as a second language programs.

The study was also limited by the unavailability of a number of urban schools to participate in the study. This limited the study by removing the possibility of comparing the low performing urban school to the low performing rural school, as well as the comparison of the high performing urban school with the high performing rural school.

### Recommendations for Further Research

One of the original intents of the researcher was to include in this study a comparison of changes in use of instructional strategies and tools, not only between low performing and high performing schools, but also between suburban/rural and urban schools. Unfortunately, the researcher was unsuccessful in attempts to gain access to the urban school district. If further research is conducted on this topic, the researcher would recommend the inclusion of this comparison of urban to suburban/rural schools. Unfortunately, the inclusion of high performing and low performing schools would most likely require the researcher to expand the study over a much larger geographical area.

Another recommendation would be to gather additional curricular information about each participating school. This might include data detailing the minimum number of graduation credits required in each of the curricula areas of English, math, science, and social studies. Another data file to include would be that of the contracted salary scales of included schools as well as the collection of information relative to the incentives to pursue educational opportunities and/or advanced degrees, offered to the teachers in the different school districts



## Conclusions

The purpose of this study was to determine (a) the differences in types of consequences realized between lower and higher performing high schools as reported by the Pennsylvania Department of Education, (b) investigating the beliefs of teachers and principals in regard to the Pennsylvania System of School Assessment (PSSA) exam, and (c) the impact of PSSA results on the instructional programs in their schools.

Available literature on high-stakes standardized testing suggests negative teacher perceptions (Moore, 1994), including dissatisfaction, a waste of valuable instructional time, and benefits not worth the cost (Hoffman et. al., 2001). Alternatively, the literature indicated that principals believed that test scores motivated educators to attempt to increase student achievement levels (McCall, 2003). That the attention to test scores helped force schools to realign the curriculum with the state standards (McCall, 2003), that principals of higher achieving schools place less pressure on teachers to increase scores on state tests while principals from lower achieving schools are deeply concerned about testing results (Reed, McDonough, Ross, & Robichaux 2001). The literature also indicated that time spent on instruction has changed and strategies have been altered since the implementation of standardized testing (Jones et al.1999; Abrams et al.2003; Vogler, 2002).

Results of this study suggest that the use of the PSSA, the high stakes test currently being administered in Pennsylvania, has definitely changed the delivery of instruction, at least in the three low performing as well as the three high performing secondary schools included in this study. The teachers included in this study reported change that is positively altering the way the curriculum is being delivered in the high

school classrooms. The use of instructional strategies has changed since the implementation of the PSSA high stakes exams, but opposite from which the literature had suggested. Teachers in all six of the participating schools have increased the use of 36 of 40 instructional strategies and tools, all considered to be consistent with best practice. Use of the individual instructional strategies and tools showed increases of 19.87% to 70%. Showing decreases in use ranging from 17% to 28%, were six items that do not properly involve students in learning. Additionally, experience in the classroom as well as the level of education pursued by the classroom teacher does have an impact on the change in the delivery of instruction. This study demonstrated that the more experienced and more educated the teacher, the more likely that the school will realize a change in the delivery of instruction. A teacher with fewer years experience and less opportunity to pursue additional education is more likely to not change the delivery of instruction.

These results lead one to conclude that the PSSA 11<sup>th</sup> grade reading and math assessments did contribute to positive changes to instructional classroom strategies as implemented by teachers.

## References

- Abrams, L. M., Pedulla, J. J., & Madaus, G. F. (2003). Views from classroom teachers' opinions of statewide testing programs. *Theory Into Practice, 42*, 18-30.
- Assessment: 2004-2005 PSSA results. (September 27, 2005). Retrieved February 28, 2006, from [http://www.pde.state.pa.us/a\\_and\\_t/cwp/view.asp?A=3&Q=115328](http://www.pde.state.pa.us/a_and_t/cwp/view.asp?A=3&Q=115328)
- Davey, L., & Monty, N. (1991). The case against a national test. *Practical Assessment, Research & Evaluation, 2(10)*. Retrieved February 27, 2007 from <http://PAREonline.net/getvn.asp?v=2&n=10>
- Delisle, J. R. (1997). How proficiency tests fall short (Let me count the ways). *Education Week, 16*, 41-42.
- Eakins, D. J., Green, D. S., & Bushell, D. (1976). The effects of an instructional test-taking unit on achievement scores. *Journal of Educational Research, 70*, 67-71.
- Edelman, J. (2001). The impact of the mandated testing program on classroom practices: Teacher perspectives. *Education, 102*, 56-59.
- Elford, G. W. (2002). *Beyond standardized testing: Better information for school accountability and management*. Lanham, MD: Scarecrow Press.
- Fink, A. (1995). *The survey kit: How to analyze survey data*. Thousand Oaks, CA: Sage Publications.
- Green, K. (Ed.). (1991). *Educational testing*. New York: Garland Publishing, Inc.
- Green, K. E., & Eicher, R. (1987). The effects of introductory tests and measurement on attitudes toward testing. *Education, 108*, 212-217.
- Haladyna, T. M. (1992). The effectiveness of several multiple-choice formats. *Applied Measurement in Education, 5*, 73-88.

- Haney, W. M., Madaus, G. F., & Lyons, R. (1993). *The fractured marketplace for standardized testing*. Boston, MA: Kluwer Academic Publishers.
- Hoffman, J. V., Assaf, L. C., & Paris, S. G. (2001). High stakes testing in reading: Today in Texas, tomorrow? *Reading Teacher*, 54, 482-493.
- Jacob, B. A. (2001). Getting tough? The impact of high school graduation exams. *Educational Evaluation and Policy Analysis*, 23, 99-121.
- Jones, M. G., Jones, B. D., Hardin, B., Chapman, L., Yarbrough, T., & Davis, M. (1999). The impact of high stakes testing on teachers and students in North Carolina. *Phi Delta Kappan*, 81, 199-204.
- Kohn, A. (1999). *The schools our children deserve: Moving beyond traditional classrooms and tougher standards*. Boston, MA: Houghton Mifflin Company.
- Kohn, A. (2000). *The case against standardized testing: Raising the scores, ruining the schools*. Portsmouth, NH: Heinemann.
- McCall, D. A. (2003). *Principals' perceptions of the MCAS: The impact of high-stakes testing in Massachusetts*. Proquest Information and Learning Company. UMI No. (3078707).
- Moon, T.R., Brighton, C.M., & Callahan, C.M. (2002). State standardized testing programs: Friend or foe of gifted education? *Roeper Review*, 25, 49-61.
- Moore, W. (1994). The devaluation of standardized testing: One district's response to a mandated assessment. *Applied Measurement in Education*, 7, 343-368.
- National Association of Elementary School Principals & National Association of Secondary School Principals (2003). *K-12 Principals guide to no child left behind*. Arlington, VA: Educational Research Service.

- National Commission on Excellence in Education (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: U. S. Department of Education. Retrieved December 3, 2004, from <http://www.ed.gov/pubs/NatAtRisk/risk.html>
- Nitko, A. (2004). *Educational Assessment of Students*. NJ: Pearson, 4th edition.
- Popham, J. W. (2001). Standardized achievement tests: Misnamed and misleading. *Education Week*, 21, 46.
- Reed, C. J., McDonough, S., Ross, M., & Robichaux, R. (2001). *Principals perceptions of high stakes testing on empowerment*. Office of Educational Research and Improvement. (ERIC Document Reproduction Service).
- Saltman, K. J., & Gabbard, D. A. (Eds). (2003). *Education as enforcement: The militarization and corporatization of schools*. NY: RoutledgeFalmer.
- Svinicki, M. & Koch, B. (1984). The great essay/multiple choice debate: Different strokes for different folks. *Innovations Abstracts*, University of Texas.
- Vogler, K. E. (2002). The impact of high stakes, state-mandated student performance assessment on teachers' instructional practices. *Education*, 123, 39-56.
- Wideen, M. F., O'Shea, T., Pye, I., & Ivany, G. (1997). High stakes testing and the teaching of science. *Canadian Journal of Education*, 22, 428-44.
- Wilde, S. (2002). *Testing and standards: A brief encyclopedia*. Portsmouth, NH: Heinemann.

## **APPENDIX A: Teacher Survey Instrument**

## Survey Instrument

### Part I

Please circle the response indicating the extent to which you have decreased or increased the use of each of the following instructional strategies since the implementation of PSSA and NCLB.

Legend:

**LD**=Large Decrease   **D**= Decrease   **S**=Same   **I**=Increase   **LI**=Large Increase  
**NA**=Not applicable

#### Instructional Strategies

1. Writing assignments	LD	D	S	I	LI	NA
2. Group projects	LD	D	S	I	LI	NA
3. Text-book based assignments	LD	D	S	I	LI	NA
4. Discussion groups	LD	D	S	I	LI	NA
5. Multiple-choice questions	LD	D	S	I	LI	NA
6. Open response questions	LD	D	S	I	LI	NA
7. True-false questions	LD	D	S	I	LI	NA
8. Use of manipulatives	LD	D	S	I	LI	NA
9. Inquiry/Investigation	LD	D	S	I	LI	NA
10. Problem-solving activities	LD	D	S	I	LI	NA
11. Work sheets	LD	D	S	I	LI	NA
12. Lesson based on current events	LD	D	S	I	LI	NA
13. Project-based assignments	LD	D	S	I	LI	NA
14. Creative/critical thinking questions	LD	D	S	I	LI	NA
15. Role playing	LD	D	S	I	LI	NA
16. Use of charts, webs, and/or outlines	LD	D	S	I	LI	NA
17. Use of response journals	LD	D	S	I	LI	NA
18. Use of portfolios	LD	D	S	I	LI	NA
19. Use of rubrics or scoring guides	LD	D	S	I	LI	NA
20. Use of exhibitions	LD	D	S	I	LI	NA
21. Interdisciplinary instruction	LD	D	S	I	LI	NA
22. Lecturing	LD	D	S	I	LI	NA
23. Modeling	LD	D	S	I	LI	NA
24. Cooperative learning/group work	LD	D	S	I	LI	NA
25. Collaborative/ team-teaching	LD	D	S	I	LI	NA
26. Peer or cross-age tutoring	LD	D	S	I	LI	NA
27. Facilitating/coaching	LD	D	S	I	LI	NA

### Instructional Materials and Tools

28. Textbooks	LD	D	S	I	LI	NA
29. Reference books	LD	D	S	I	LI	NA
30. Supplementary books	LD	D	S	I	LI	NA
31. Primary source material	LD	D	S	I	LI	NA
32. Newspaper/magazines	LD	D	S	I	LI	NA
33. Audiovisual materials	LD	D	S	I	LI	NA
34. Lab equipment	LD	D	S	I	LI	NA
35. Calculators	LD	D	S	I	LI	NA
36. Computers/educational software	LD	D	S	I	LI	NA
37. Computers/internet and/or on-line research service	LD	D	S	I	LI	NA
38. Manipulatives	LD	D	S	I	LI	NA
39. Maps/globes/atlasses	LD	D	S	I	LI	NA
40. Visual aids (e.g. posters, graphs)	LD	D	S	I	LI	NA

### Part II

Please circle the number indicating your responses to the statements below.

Legend:

**SD**=Strongly Disagree    **D**=Disagree    **U**=Undecided    **A**=Agree    **SA**=Strongly Agree

The following has influenced changes in my instructional practices since the implementation of PSSA and NCLB.

41. Personal desire to make changes	SD	D	U	A	SA
42. Belief that such changes will benefit students	SD	D	U	A	SA
43. Changes in the types assessment used for school accountability	SD	D	U	A	SA
44. Interest in helping my school improve PSSA scores	SD	D	U	A	SA
45. Interest in helping my students attain PSSA scores that will allow them to graduate	SD	D	U	A	SA
46. Interest in avoiding sanctions at my school	SD	D	U	A	SA
47. Interactions with school principal(s)	SD	D	U	A	SA
48. Interactions with colleagues	SD	D	U	A	SA
49. Staff development in which I have participated	SD	D	U	A	SA
50. Interactions with parents	SD	D	U	A	SA
51. Curriculum was aligned to coordinate with state standards	SD	D	U	A	SA



### Part III

Please circle the responses that describe you.

52. Male    Female

53. **Teaching Experience**

- 3 years or less
- 4-12 years
- 13-19 years
- 20-27 years
- 28 or more years

54. **Education** (Please circle the highest level obtained)

- Bachelor Degree
- Bachelor + 15
- Masters
- Masters +15
- Masters +30
- Masters +45
- Masters +60
- Doctorate

55. **Teaching Assignment** (Please indicate your teaching assignment)

- English
- Mathematics
- Science
- Social Studies
- Art
- Music
- Physical Education
- Business
- Technology Education
- Special Education

56. How has PSSA testing impacted the curriculum in your specific content area?

## **APPENDIX B: Principal Interview Questions**

## Part I

Schools across the state of Pennsylvania have been participating in the administration of the PSSA exams for a number of years. Since the implementation of No Child Left Behind and with a goal of 100% proficiency by the year 2014, a number of schools have been placed under pressure to drastically improve student performance.

1. Please discuss your school in relation to the PSSA test results and to what extent has your school been impacted by the results.
2. As the results of PSSA have been reported, and as the instructional leader of the school, have you exerted more pressure
  - a. on the teachers?
  - b. on the students?
3. As the results of PSSA have been reported, and as the instructional leader of the school, have you had more pressure exerted on you from central administration?
4. What changes in curriculum has occurred as a consequence of the PSSA results?
5. Have changes occurred in the area of teaching strategies by teachers in your school as a consequence of PSSA test results?
6. As a consequence of PSSA results, what particular teaching strategies are teachers using
  - a. more frequently?
  - b. Less frequently?
7. Have changes occurred in the use of materials in your schools as a consequence of the PSSA results?
8. As a consequence of PSSA results, what particular materials are used
  - a. more frequently?
  - b. less frequently?
9. What are some of the influences that caused the changes to curriculum, teaching strategies, or use of materials in the school?

## Part II

### Demographic Data

10. Male or Female
11. Number of Years Teaching Experience?
12. Number of Years in Administration?
13. Number of years in this school as teacher \_\_\_\_\_, as principal \_\_\_\_\_?
14. Level of Education
  - a. Master Degree
  - b. Doctorate
15. Area of teaching certification?
16. Number of years at current position?
17. Number of years at current school?

## **APPENDIX C: Permission**

**From:** "Kenneth Vogler" <kvogler@gwm.sc.edu>  
**To:** <SSignorino@fairview.iu5.org>  
**Date:** 2/23/06 8:44PM  
**Subject:** Re: Request

Sam:

You have my permission to use the survey instrument. Good luck with your study.

Ken

>>> "Sam Signorino" <SSignorino@fairview.iu5.org> 02/23/06 7:27 AM >>>

Dr. Vogler,

I am a doctoral candidate at Duquesne University as well as a high school principal in Pennsylvania. I am submitting a dissertation proposal that would permit research concerning the consequences of the PA state test (PSSA). My study titled "DIFFERENCES IN TYPES AND SEVERITY OF CONSEQUENCES OF HIGH-STAKES TESTING AS REALIZED BETWEEN LOW AND HIGH PERFORMING HIGH SCHOOLS AND THE IMPACT ON THE INSTRUCTIONAL PROGRAMS IN THEIR SCHOOLS" will be looking at how teachers and school districts are changing/eliminating/altering their teaching practices and strategies as a result of the Pennsylvania System of School Assessment (PSSA) scores as received from the Department of Education.

If you are the same Ken Vogler, I read your article--THE IMPACT OF HIGH-STAKES, STATE-MANDATED STUDENT PERFORMANCE ASSESSMENT ON TEACHER'S INSTRUCTIONAL PRACTICES, By: Vogler, Kenneth E., Education, 00131172, Fall 2002, Vol. 123, Issue 1

I also spent some time looking at the survey that was used to collect data for the study in Massachusetts.

My study will involve high school teachers in Northwestern PA from a small number of high performing schools and low performing schools. I hope to see that the high performing schools are not abandoning good teaching practices and strategies as the low performing schools may be doing.

With your permission, I would like to use your survey from the above-mentioned study. There would be slight alterations from the form I found in the publication. These changes would be minor. For example, the directions would be changed to include PSSA instead of the MCAS, and in the demographics (part III) I will expand the area of teaching assignment to include Social Studies, Art, Music, Physical Education, Business, and Technology Education.

The only other change I am considering is also in part III, levels of education.

Please let me know if this request is something you might consider. If further information is necessary please feel free to contact me.

Thank you for your consideration!!  
Samuel Signorino  
ssignorino@fairview.iu5.org  
814-474-2600 ext. 3101 (office)

Fairview High School  
7460 McCray Road  
Fairview, PA 16415